



Best Available Copy

Atty. Dkt. No. 039386-2282

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

Applicants: Policky et al.


Title: HUMAN CYSTEINYL
LEUKOTRIENE RECEPTORS

Appl. No.: 09/980,049

Filing Date: 11/28/2001

Examiner: Ulm, John D.

Art Unit: 1649

CERTIFICATE OF EXPRESS MAILING	
I hereby certify that this correspondence is being deposited with the United States Postal Service's "Express Mail Post Office To Addressee" service under 37 C.F.R. § 1.10 on the date indicated below and is addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.	
EV 593167980 US (Express Mail Label Number)	April 21, 2006 (Date of Deposit)
Deborah A. Kocorowski (Printed Name)	
 (Signature)	

Mail Stop **APPEAL BRIEF - PATENTS**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL

Transmitted herewith please find the following documents for the above-identified patent application.

- [X] Brief on Appeal (17 pages).
- [X] Exhibit 1 (59 pages).
- [X] Exhibit 2 (84 pages).
- [X] Exhibit 3 (6 pages).
- [X] Credit Card Payment Form for \$500.00 (1 page).

The Commissioner is hereby authorized to charge any additional fees which may be required regarding this application under 37 C.F.R. §§ 1.16-1.17, or credit any overpayment, to Deposit Account No. 19-0741. Should no proper payment be enclosed herewith, as by the credit card payment form being unsigned, providing incorrect information resulting in a rejected credit card transaction, or even entirely missing, the Commissioner is authorized to charge the unpaid amount to Deposit Account No. 19-0741.

If any extensions of time are needed for timely acceptance of papers submitted herewith, the Applicants hereby petition for such extension under 37 C.F.R. § 1.136 and authorize payment of any such extensions fees to Deposit Account No. 19-0741.

Respectfully submitted,

Date

April 21, 2006

By

M. Scott McBride

FOLEY & LARDNER LLP
Customer Number: 22428

M. Scott McBride
Attorney for the Applicants
Registration No. 52,008
Telephone: (414) 297-5529
Facsimile: (414) 297-4900



Atty. Dkt. No. 039386-2282

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

Applicants: Policky et al.

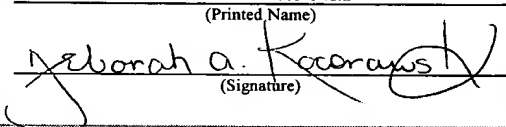
Title: HUMAN CYSTEINYL
LEUKOTRIENE RECEPTORS

Appl. No.: 09/980,049

Filing Date: 11/28/2001

Examiner: Ulm, John D.

Art Unit: 1649

CERTIFICATE OF EXPRESS MAILING	
I hereby certify that this correspondence is being deposited with the United States Postal Service's "Express Mail Post Office To Addressee" service under 37 C.F.R. § 1.10 on the date indicated below and is addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.	
EV 593167980 US	April 21, 2006
(Express Mail Label Number)	(Date of Deposit)
Deborah A. Kocorowski	
(Printed Name)	
	
(Signature)	

Mail Stop **APPEAL BRIEF - PATENTS**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

BRIEF ON APPEAL

Under the provisions of 37 C.F.R. § 41.37, this Appeal Brief is being filed together with a Credit Card Payment Form for the amount of **\$500.00** covering the appeal fee as provided under 37 C.F.R. § 41.20(b)(2). If this fee is deemed to be insufficient, authorization is hereby given to charge any deficiency (or credit any balance) to the undersigned deposit account 19-0741. This brief is timely filed.

04/25/2006 SFELEKE1 00000054 09980049

01 FC:1402

500.00 DP

I. REAL PARTY IN INTEREST

The real party in interest is INCYTE CORPORATION.

II. RELATED APPEALS AND INTERFERENCES

The Appellants are unaware of any related appeals or interferences.

III. STATUS OF AMENDMENTS

No amendment to the claims has been filed subsequent to the objection to and rejection of claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 in the final Office Action dated September 22, 2005. The amendment filed on July 13, 2005 is objected to under 35 U.S.C. § 132(a) for allegedly introducing new matter.

IV. STATUS OF CLAIMS

Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand objected to under 35 U.S.C. § 132 in the final Office Action mailed on September 22, 2005. Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 are currently pending and stand rejected under 35 U.S.C. §§ 101, 112 and 102(a) in the final Office Action mailed on September 22, 2005. The objection to and rejection of claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 is appealed.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The claimed subject matter relates to nucleic acids and polypeptides of G-protein coupled receptors, specifically G-protein coupled receptors with cysteinyl leukotriene receptor activity. The claimed subject matter also relates to compositions that include the polypeptides and methods of using the polypeptides to identify compounds, such as compounds that modulate activity of the polypeptides.

Claim 1 recites "An isolated polypeptide" that may include "a polypeptide comprising an amino acid sequence of SEQ ID NO:1." The recited polypeptide may have cysteinyl leukotriene receptor activity. Claim 11 recites "An isolated polynucleotide" that may include "a polynucleotide comprising a polynucleotide sequence of SEQ ID NO:7."

At paragraph [0004] the specification states that “GCPRs [G-coupled protein receptors] include receptors for...lipid mediators of inflammation (*e.g.*, ...*leukotrienes*).” (See U.S. 2004-0220092, paragraph [0004]). Additionally, the specification at paragraph [0130] states that “Tables 2 and 3 *summarize the properties of the polypeptides of the invention.*” (See *id.*, paragraph [0130]). Table 2 provides that the nearest Genbank homolog to SEQ ID NO: 1, a polypeptide of the invention, is a *cysteinyl leukotriene receptor*, a member of the G-protein coupled receptor family. (See paragraph [0339], Table 2).

VI. GROUND OF REJECTION TO BE REVIEWED ON APPEAL

The amendment filed on July 13, 2005 stands objected to under 35 U.S.C. § 132(a) for allegedly introducing “new matter,” including the Title of the application, and the recitation of “wherein the polypeptide has leukotriene receptor activity,” in claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61. Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 101 for lack of specific, substantial and credible utility. Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 112, first paragraph, for lack of enabling description and lack of written description. Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 102(a) as being allegedly anticipated by Takasaki *et al.*, B.B.R.C. 274(2):316-322, (Aug. 2000) (“Takasaki”).

VII. ARGUMENT

All of the objections and rejections are based on the contention that the present application and its priority application (U.S. provisional application no. 60/199,084) do not explicitly disclose that the polypeptide of SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Accordingly, the arguments presented in Section VII.A. below demonstrate that the specification indeed does disclose that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. The arguments in Sections VII.B.-F. address the specific objections and rejections put forth in the Final Office Action mailed on September 22, 2005, and establish that each objection and rejection should be reversed.

A. Assertion that SEQ ID NO:1 is a Human G-Coupled Protein Receptor for Cysteinyl Leukotrienes

The present application generally relates to G-Protein Coupled Receptors (*i.e.*, “GPCRs” or “GCRECs”). In fact, the application as filed was entitled “G-Protein Coupled Receptors.” In the specification, the polypeptide of SEQ ID NO:1 is referred to as “G-Protein Coupled Receptor 1” or “GCREC-1.” At paragraph [0130], the specification states that “[t]ogether, Tables 2 and 3 summarize the properties of polypeptides of the invention, and *these properties establish that the claimed polypeptides are G-protein coupled receptors.*” (*See id.*, (emphasis added)). Therefore, it is undisputable that SEQ ID NO:1 is disclosed to be a human G-Protein Coupled Receptor. Because SEQ ID NO:1 is disclosed to be a human G-Protein Coupled Receptor, SEQ ID NO:1, like any receptor, must have a specific ligand.

The specification indicates that the polypeptide of SEQ ID NO:1 is a G-Coupled Protein Receptor for cysteinyl leukotrienes. At paragraph [0004], the specification states that “GPCRs include receptors for...lipid mediators of inflammation (*e.g.*,...*leukotrienes*)” (*See id.*, paragraph [0004]; (emphasis added)). The specification states that Table 2 “summarizes the properties of polypeptides of the invention.” (*See id.*, paragraph [0130]). Table 2 provides the following:

TABLE 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	5628963CD1	g10442008	0	Cysteinyl leukotriene receptor CYSLT2 [<i>Homo sapiens</i>] (Heise, C. E. et al. (2000) J. Biol. Chem. 275: 30531–30536)

(*See id.*, paragraph [0339], Table 2).

At Table 2, the specification presents “Cysteinyl leukotriene receptor CYSLT2 [Homo sapiens]” as the nearest GenBank homolog for the polypeptide of SEQ ID NO:1 and indicates that there is **a zero (0) probability score** for a match by BLAST analysis between “Cysteinyl leukotriene receptor CYSLT2 [Homo sapiens]” and the polypeptide of SEQ ID NO:1. (*See id.*). In Table 2, a low probability score is indicative of a low probability of having obtained a match by chance, and accordingly, a probability score of zero (0) means that there is zero probability of having obtained the match by chance. At Table 3, the specification indicates that SEQ ID NO:1 includes “Signature Sequences, Domains and Motifs,” which are characteristic of cysteinyl leukotriene receptors and supports the assertion that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. (*See id.*, paragraph [0340]).

Accordingly, the only reasonable interpretation of the disclosure in the specification, reading together paragraph [0004] (stating that “GPCRs include receptors for...*leukotrienes*); paragraph [0130] (stating that “[t]ogether, Tables 2 and 3 summarize the properties of polypeptides of the invention); and the data in Tables 2 and 3 (indicating, *inter alia*, that there is a zero probability of matching SEQ ID NO:1 and Cysteinyl leukotriene receptor CYSLT2 by chance); is that SEQ ID NO:1 is asserted to be a human G-Protein Coupled Receptor for cysteinyl leukotrienes.

Provisional Application No. 60/199,084 (“the ‘084 Application”) includes the same assertion. At page 5, lines 1-2, the ‘084 Application discloses that SEQ ID NO:1 is referred to as “G-Protein Coupled Receptor 1” or “GCREC-1.” At page 2, lines 4-10, the ‘084 Application states that “GPCRs include receptors for...lipid mediators of inflammation (e.g.,...*leukotrienes*).” Table 2 of the ‘084 Application provides:

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	5628963CD1	g5353887	2.2e-54	Cysteinyl leukotriene LTD4 receptor [Homo sapiens]
2	1453124CD1	g5525078	1.0e-134	Seven transmembrane receptor [Rattus norvegicus]

Table 2 shows that “Cysteinyl leukotriene LTD4 receptor [Homo sapiens]” is the nearest GenBank homolog for the polypeptide of SEQ ID NO:1 and indicates that there is very low probability score for a match by BLAST analysis between “Cysteinyl leukotriene LTD4 receptor [Homo sapiens]” and the polypeptide of SEQ ID NO:1 (*i.e.*, a 2.2×10^{-54} probability of having obtained the match by chance). At Table 3, the specification of the ‘084 Application indicates that SEQ ID NO:1 includes “Signature Sequences, Domains and Motifs,” which are characteristic of cysteinyl leukotriene receptors and supports the assertion that SEQ ID NO:1 is a human G-Coupled Protein Receptor for cysteinyl leukotrienes.

Therefore, the specifications of the present application and the priority ‘084 Application include assertions that the polypeptide of SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Whether the asserted utility is credible, specific, and substantial is a *separate issue* as discussed below.

B. Objection – 35 U.S.C. § 132, “New Matter”

The amendment filed on July 13, 2005, was objected to for allegedly introducing new matter into the disclosure. In particular, the Title, which recites “HUMAN CYSTEINYL LEUKOTRIENE RECEPTORS,” and the claim limitation “wherein the polypeptide has cysteinyl leukotriene receptor activity” are alleged to be “new matter completely without support in the specification.” (*See* Office Action dated September 22, 2005, page 2).

As indicated above, the present specification and the priority ‘084 Application assert that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Therefore, recitation of “HUMAN CYSTEINYL LEUKOTRIENE RECEPTORS,” and “wherein the polypeptide has cysteinyl leukotriene receptor activity” are not “new matter.”

The objection under 35 U.S.C. § 132, for alleged “new matter” should be reversed.

C. Rejection – 35 U.S.C. § 101, “Utility”

Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 101 allegedly “because they are drawn to an invention with no apparent or disclosed specific and substantial credible utility.” The Appellants respectfully disagree.

The utility requirement may be viewed as requiring: (1) that an application include an asserted utility for the claimed subject matter; and further, (2) that the asserted utility be a specific and substantial credible utility. As indicated above and contrary to the Office's position, the present specification and the '084 Application assert that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Whether this asserted utility is "specific," "substantial," and "credible" are *separate issues* that are addressed below.

As indicated above, it is the Appellants' position that the specification does assert that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. However, as stated in the Advisory Action dated January 27, 2006, it is the Office's position that the present specification does not assert that SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes because "no conclusions are provided with respect to the functions of a receptor protein comprising SEQ ID NO:1." The Appellants disagree. It is respectfully noted that an "assertion" may be defined as "something asserted positively, often with no support or attempt of proof." (*See, e.g., American Heritage Dictionary, definition of "assertion."*) The Office's position requires more than an "assertion."

Further, the Advisory Action dated January 27, 2006 indicates that the specification does include "assertions" with respect to utility and function for other SEQ ID NOs. As stated in the Advisory Action:

The text at the top page 26 of the instant specification provides an analysis of the data allegedly presented in Tables 2 and 3 with regard to SEQ ID NO:3, which is the only sequence from Table 2 whose function is [discussed] therein. That text observes that the closest [GenBank] homologs to SEQ ID NO:3 are known serotonin (5HT) receptors. However, the specification does not assert that SEQ ID NO:3 is the amino acid sequence of a serotonin receptor. Instead, *it asserts that SEQ ID NO:3 is the amino acid [sequence] of a "G-protein coupled receptor"* that "acts as a receptor for serotonin, another neurotransmitter molecule, or another [biogenic] amine." This hardly rises to the level of an assertion that a specific sequence presented in Table 2 has the same function as the closest known [GenBank] homolog and certainly does not constitute a disclosure of a specific and substantial utility for a protein comprising any one of those sequences.

(See Advisory Action dated January 27, 2006, page 3).

In regard to these statements in the Advisory Action, it is unclear to the Appellants how disclosure in the specification related to the asserted utility of SEQ ID NO:3 can render insufficient an asserted utility with respect to SEQ ID NO:1. However, the Appellants respectfully contend that the Office's position as stated in the Advisory Action is inconsistent with the maintained rejection. The Advisory Action states that the text of the specification, describing results presented in Tables 2 and 3, "**asserts** that SEQ ID NO:3 is the amino acid [sequence] of a '**G-protein coupled receptor**.'" (See *id.*). With respect to SEQ ID NO:3, Table 2 presents "**G protein-coupled seven-transmembrane receptor [Oryzias latipes]**" as the nearest GenBank homolog and indicates that there is a 6.10×10^{-82} probability score for a match by BLAST analysis between "**G protein-coupled seven-transmembrane receptor [Oryzias latipes]**" and SEQ ID NO:3. (See U.S. 2004-0022092, paragraph [0339], Table 2, (emphasis added)). In discussing the results presented in Tables 2 and 3, the specification further states that "Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide **further** corroborative evidence that **SEQ ID NO:3 is a G-protein coupled receptor**." (See *id.*, paragraph [0130], (emphasis added), notably referring to the "Data from BLIMPS, MOTIFS, and PROFILESCAN analyses" as "**further** corroborative evidence"). As such, the text of the present specification indicates, as recognized in the Advisory Action, that a presentation in Table 2 **is an assertion** that the polypeptide of SEQ ID NO:3 **is a G protein-coupled receptor**. The Appellants request that the Office apply the same standard with respect to the presentation in Table 2 that relates to SEQ ID NO:1 and "Cysteinyl leukotriene receptor CYSLT2 [Homo sapiens]."

As is well appreciated, post-filing evidence may be used to substantiate an asserted utility. The asserted utility of SEQ ID NO:1 as a human G-Coupled Protein Receptor for cysteinyl leukotrienes, otherwise referred to as "human cysteinyl leukotriene 2 receptor," has been confirmed, for example by Heise *et al.*, Characterization of the Human Cysteinyl Leukotriene 2 Receptor, J. BIOL. CHEM. (September 2000), Vol. 275, No. 39, pp 30531-30536, [hereinafter "Heise"]. Therefore, the claimed subject matter has a "specific" utility.

Furthermore, Heise indicates that cysteinyl leukotriene receptors are known to mediate contractile and inflammatory actions of cysteinyl leukotriene ligands. In particular, cysteinyl leukotriene receptors are known to mediate bronchoconstrictive and inflammatory actions in the lungs. Therefore, the claimed subject matter has a “substantial” utility.

Finally, because others have tested and confirmed the asserted utility for SEQ ID NO:1, the claimed subject matter has a “credible” utility.

For these reasons, the rejection under 35 U.S.C. § 101 for alleged lack of utility should be reversed.

D. Rejection – 35 U.S.C. § 112, first paragraph, “Enablement”

Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 112, first paragraph, allegedly “as failing to adequately teach how to use the invention for those reasons given above with regard to the rejection of these claims under 35 U.S.C. § 101.” For the reasons stated above with respect to the rejection under 35 U.S.C. § 101, the rejection should be reversed.

E. Rejection – 35 U.S.C. § 112, first paragraph, “Written Description”

Claims 1-7, 9, 11, 16, 17, 19, 22, 26, and 57-61 stand rejected under 35 U.S.C. § 112, first paragraph, allegedly “as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.” In particular, the Office’s position is that the application does not “provide[] a written description of a polypeptide having the amino acid sequence presented in SEQ ID NO:1 and ‘cysteinyl leukotriene activity.’” Further, the Office’s position is that “the phrase ‘cysteinyl leukotriene receptor activity’ is without support in the instant application and a relationship between this activity and SEQ ID NO:1 is a new inventive concept.”

As indicated above, the present specification and the priority '084 Application assert that the polypeptide of SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Therefore, a polypeptide of SEQ ID NO:1 having cysteinyl leukotriene receptor activity is disclosed in the present specification and the '084 Application and is not a new inventive concept without support in the specification.

The rejection under 35 U.S.C. § 112, first paragraph, for inadequate "written description" should be reversed.

F. Rejection – 35 U.S.C. § 102, "Takasaki et al."

Claims 1-7, 9, 11, 16, 17, 19, 22, 26 and 57-61 stand rejected under 35 U.S.C. § 102(a) as being allegedly anticipated by Takasaki *et al.*, B.B.R.C. 274(2):316-322, (Aug. 2000)" ("Takasaki").

As indicated above, the present application claims the benefit of U.S. provisional Application No. 60/199,084, filed on April 20, 2000. Also as indicated above, the '084 Application includes substantially similar disclosure as the present application with respect to the polypeptide of SEQ ID NO:1. In particular, the '084 Application includes an assertion that the polypeptide of SEQ ID NO:1 is a human G-Protein Coupled Receptor for cysteinyl leukotrienes. Therefore, the '084 Application fully supports the presently claimed subject matter under 35 U.S.C. § 112, and the pending claims are entitled to a priority date at least as of the filing date of the '084 Application (*i.e.*, April 20, 2000), which predates the Takasaki publication.

For these reasons, the rejection under 35 U.S.C. § 102(a) should be reversed.

VIII. SUMMARY

The Appellants request that the Honorable Board reverse the outstanding final rejections of and objections to the claims.

Respectfully submitted,

Date April 21, 2006

FOLEY & LARDNER LLP
Customer Number: 22428

By M. Scott McBride

M. Scott McBride
Attorney for the Applicants
Registration No. 52,008
Telephone: (414) 297-5529
Facsimile: (414) 297-4900

CLAIMS APPENDIX

1. (Previously presented) An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence of SEQ ID NO:1, and
 - b) a polypeptide comprising an amino acid sequence having at least about 95% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.
2. (Previously presented) The isolated polypeptide of claim 1 comprising an amino acid sequence of SEQ ID NO:1.
3. (Previously presented) An isolated polynucleotide encoding a polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence of SEQ ID NO:1; and
 - b) a polypeptide comprising an amino acid sequence having at least about 95% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.
4. (Previously presented) The isolated polynucleotide of claim 3 encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:1.
5. (Previously presented) The isolated polynucleotide of claim 4 comprising a polynucleotide sequence of SEQ ID NO:7.
6. (Previously presented) A recombinant polynucleotide comprising a promoter sequence operably linked to the polynucleotide of claim 3.
7. (Previously presented) An isolated cell transformed with the recombinant polynucleotide of claim 6.
8. (Cancelled).

9. (Previously presented) A method for producing the polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. (Cancelled).

11. (Previously presented) An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence of SEQ ID NO:7,

b) a polynucleotide comprising a polynucleotide sequence having at least about 95% sequence identity to a polynucleotide sequence of SEQ ID NO:7, wherein the polynucleotide encodes a polypeptide that has cysteinyl leukotriene receptor activity,

c) a polynucleotide complementary to the polynucleotide of a),

d) a polynucleotide complementary to the polynucleotide of b), and

e) an RNA equivalent of a)-d).

12.-15. (Cancelled).

16. (Previously presented) A composition comprising the polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. (Previously presented) The composition of claim 16, wherein the polypeptide has an amino acid sequence of SEQ ID NO:1.

18. (Cancelled).

19. (Previously presented) A method for screening a compound for effectiveness as an agonist of the polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising the polypeptide of claim 1 to a compound,
- and
- b) detecting agonist activity in the sample.

20.-21. (Cancelled).

22. (Previously presented) A method for screening a compound for effectiveness as an antagonist of the polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising the polypeptide of claim 1 to a compound,
- and
- b) detecting antagonist activity in the sample.

23.-25. (Cancelled).

26. (Original) A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27.-56. (Cancelled).

57. (Previously presented) The isolated polypeptide of claim 1 comprising an amino acid sequence having at least about 96% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.

58. (Previously presented) The isolated polypeptide of claim 1 comprising an amino acid sequence having at least about 97% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.

59. (Previously presented) The isolated polypeptide of claim 1 comprising an amino acid sequence having at least about 98% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.

60. (Previously presented) The isolated polypeptide of claim 1 comprising an amino acid sequence having at least about 99% sequence identity to an amino acid sequence of SEQ ID NO:1, wherein the polypeptide has cysteinyl leukotriene receptor activity.

61. (Previously presented) The isolated polypeptide of claim 1, wherein the polypeptide is a human polypeptide.

EVIDENCE EXHIBIT

In support of this Brief on Appeal, the Appellants additionally append copies of the following evidence that was introduced and entered into the record during prosecution:

- Exhibit 1. US 2004/00220092 A1 (published specification of the present application)
- Exhibit 2. U.S. Provisional Application No. 60/199,084
- Exhibit 3. Heise *et al.*, Characterization of the Human Cysteinyl Leukotriene 2 Receptor, J. BIOL. CHEM. (September 2000), Vol. 275, No. 39, pp 30531-30536

RELATED PROCEEDINGS APPENDIX

The Applicants have no information regarding related proceedings to submit with this appeal brief.

04/20/00



JVS09 U.S. PTO

04-24-00

A/P/10V
Seq Box

Atty. Docket No.: PI-0072 P

Type a plus sign (+) inside this box

+

"Express Mail" mailing label number E1582 175 044 US. I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10 on the date indicated and is addressed to: Assistant Commissioner for Patents, Box Provisional Patent Application, Washington, D.C. 20231 on 4-20-00.

By:

Nancy Ramos Printed: Nancy Ramos

PROVISIONAL APPLICATION COVER SHEET (Large Entity)

JCS41 U.S. PTO
 60/199084
 04/20/00

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2).

INVENTOR(S)/APPLICANT(S)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Policky	Jennifer	L.	San Jose, California
Tribouley	Catherine	M.	San Francisco, California
Tang	Y. Tom		San Jose, California
Baughn	Mariah	R.	San Leandro, California
Graul	Rick		San Francisco, California
Khan	Farrah	A.	Mountain View, California
Nguyen	Danniel	B.	San Jose, California
Patterson	Chandra		Menlo Park, California
TITLE OF THE INVENTION (280 characters max)			
G-PROTEIN COUPLED RECEPTORS			
CORRESPONDENCE			
INCYTE GENOMICS, INC. Patent Department 3160 Porter Drive Palo Alto, California 94304		Phone: (650) 855-0555 Fax: (650) 849-8886 or (650) 845-4166	

57228

1

EXHIBIT

2

Blumberg Inc. 5118

Atty. Docket No.: PI-0072 P	Type a plus sign (+) inside this box ➡	+
-----------------------------	---	---

ENCLOSED ARE:	
1.	Return postcard;
2.	Provisional Application Cover Sheet w/certificate of Express Mailing (in duplicate, 2 pages);
3.	Submission under 37 CFR § 1.821-1.825 Sequence Listing Statement with one Computer-Readable Diskette attached;
4.	<u>62</u> Pages of Specification (1-62);
5.	<u>3</u> Pages of Claims (63-65);
6.	<u>1</u> Pages of Abstract (66);
7.	<u>10</u> Pages of Tables (Tables 1-7);
8.	<u>6</u> Pages of Sequence listing (1-6).
METHOD OF PAYMENT	
The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number: <u>09-0108</u>	
PROVISIONAL FILING FEE AMOUNT: \$150.00	

The invention was not made by an agency of the United States Government or under a contract with an agency of the United States Government.

Date:

20 April 2000.

Respectfully submitted,

INCYTE PHARMACEUTICALS, INC.

Adam Warwick Bell, D.Phil.

Reg. No. 43,490

Direct Dial Telephone: (650) 621-7542

G-PROTEIN COUPLED RECEPTORS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of G-protein coupled receptors and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections.

BACKGROUND OF THE INVENTION

Signal transduction is the general process by which cells respond to extracellular signals. Signal transduction across the plasma membrane begins with the binding of a signal molecule, e.g., a hormone, neurotransmitter, or growth factor, to a cell membrane receptor. The receptor, thus activated, triggers an intracellular biochemical cascade that ends with the activation of an intracellular target molecule, such as a transcription factor. This process of signal transduction regulates all types of cell functions including cell proliferation, differentiation, and gene transcription. The G-protein coupled receptors (GPCRs), encoded by one of the largest families of genes yet identified, play a central role in the transduction of extracellular signals across the plasma membrane. GPCRs have a proven history of being successful therapeutic targets.

GPCRs are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which together form a bundle of antiparallel alpha (α) helices. GPCRs range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) Eur. J. Biochem. 196:1-10; Coughlin, S.R. (1994) Curr. Opin. Cell Biol. 6:191-197). The amino-terminus of a GPCR is extracellular, is of variable length, and is often glycosylated. The carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops alternate with intracellular loops and link the transmembrane domains. Cysteine disulfide bridges linking the second and third extracellular loops may interact with agonists and antagonists. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account, in part, for structural and functional features of the receptor. In most cases, the bundle of α helices forms a ligand-binding pocket. The extracellular N-terminal segment, or one or more of the three extracellular loops, may also participate in ligand binding. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. In turn, the large, third intracellular loop of the activated receptor interacts with a heterotrimeric guanine nucleotide binding (G) protein complex which mediates further intracellular signaling activities, including the activation of second messengers such as cyclic AMP (cAMP), phospholipase C, and inositol triphosphate, and the interaction of the activated GPCR with ion channel proteins. (See, e.g., Watson, S. and S.

Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 2-6; Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego CA, pp. 162-176; Baldwin, J.M. (1994) *Curr. Opin. Cell Biol.* 6:180-190.)

GPCRs include receptors for sensory signal mediators (e.g., light and olfactory stimulatory molecules); adenosine, γ -aminobutyric acid (GABA), hepatocyte growth factor, melanocortins, neuropeptide Y, opioid peptides, opsins, somatostatin, tachykinins, vasoactive intestinal polypeptide family, and vasopressin; biogenic amines (e.g., dopamine, epinephrine and norepinephrine, histamine, glutamate (metabotropic effect), acetylcholine (muscarinic effect), and serotonin); chemokines; lipid mediators of inflammation (e.g., prostaglandins and prostanoids, platelet activating factor, and leukotrienes); and peptide hormones (e.g., bombesin, bradykinin, calcitonin, C5a anaphylatoxin, endothelin, follicle-stimulating hormone (FSH), gonadotropic-releasing hormone (GnRH), neurokinin, and thyrotropin-releasing hormone (TRH), and oxytocin). GPCRs which act as receptors for stimuli that have yet to be identified are known as orphan receptors.

The diversity of the GPCR family is further increased by alternative splicing. Many GPCR genes contain introns, and there are currently over 30 such receptors for which splice variants have been identified. The largest number of variations are at the protein C-terminus. N-terminal and cytoplasmic loop variants are also frequent, while variants in the extracellular loops or transmembrane domains are less common. Some receptors have more than one site at which variance can occur. The splicing variants appear to be functionally distinct, based upon observed differences in distribution, signaling, coupling, regulation, and ligand binding profiles (Kilpatrick, G.J. et al. (1999) *Trends Pharmacol. Sci.* 20:294-301).

GPCRs can be divided into three major subfamilies: the rhodopsin-like, secretin-like, and metabotropic glutamate receptor subfamilies. Members of these GPCR subfamilies share similar functions and the characteristic seven transmembrane structure, but have divergent amino acid sequences. The largest family consists of the rhodopsin-like GPCRs, which transmit diverse extracellular signals including hormones, neurotransmitters, and light. Rhodopsin is a photosensitive GPCR found in animal retinas. In vertebrates, rhodopsin molecules are embedded in membranous stacks found in photoreceptor (rod) cells. Each rhodopsin molecule responds to a photon of light by triggering a decrease in cGMP levels which leads to the closure of plasma membrane sodium channels. In this manner, a visual signal is converted to a neural impulse. Other rhodopsin-like GPCRs are directly involved in responding to neurotransmitters. These GPCRs include the receptors for adrenaline (adrenergic receptors), acetylcholine (muscarinic receptors), adenosine, galanin, and glutamate (N-methyl-D-aspartate/NMDA receptors). (Reviewed in Watson, S. and S. Arkininstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 7-9, 19-22, 32-35, 130-131, 214-216, 221-222; Habert-Ortoli, E. et al. (1994) *Proc. Natl. Acad. Sci. USA*

91:9780-9783.)

The galanin receptors mediate the activity of the neuroendocrine peptide galanin, which inhibits secretion of insulin, acetylcholine, serotonin and noradrenaline, and stimulates prolactin and growth hormone release. Galanin receptors are involved in feeding disorders, pain, depression, and Alzheimer's disease (Kask, K. et al. (1997) Life Sci. 60:1523-1533). Other nervous system rhodopsin-like GPCRs include a growing family of receptors for lysophosphatidic acid and other lysophospholipids, which appear to have roles in development and neuropathology (Chun, J. et al. (1999) Cell Biochem. Biophys. 30:213-242).

The largest subfamily of GPCRs, the olfactory receptors, are also members of the rhodopsin-like GPCR family. These receptors function by transducing odorant signals. Numerous distinct olfactory receptors are required to distinguish different odors. Each olfactory sensory neuron expresses only one type of olfactory receptor, and distinct spatial zones of neurons expressing distinct receptors are found in nasal passages. For example, the RA1c receptor which was isolated from a rat brain library, has been shown to be limited in expression to very distinct regions of the brain and a defined zone of the olfactory epithelium (Raming, K. et al. (1998) Receptors Channels 6:141-151). However, the expression of olfactory-like receptors is not confined to olfactory tissues. For example, three rat genes encoding olfactory-like receptors having typical GPCR characteristics showed expression patterns not only in taste and olfactory tissue, but also in male reproductive tissue (Thomas, M.B. et al. (1996) Gene 178:1-5).

Members of the secretin-like GPCR subfamily have as their ligands peptide hormones such as secretin, calcitonin, glucagon, growth hormone-releasing hormone, parathyroid hormone, and vasoactive intestinal peptide. For example, the secretin receptor responds to secretin, a peptide hormone that stimulates the secretion of enzymes and ions in the pancreas and small intestine (Watson, *supra*, pp. 278-283). Secretin receptors are about 450 amino acids in length and are found in the plasma membrane of gastrointestinal cells. Binding of secretin to its receptor stimulates the production of cAMP.

Examples of secretin-like GPCRs implicated in inflammation and the immune response include the EGF module-containing, mucin-like hormone receptor (Emr1) and CD97 receptor proteins. These GPCRs are members of the recently characterized EGF-TM7 receptors subfamily. These seven transmembrane hormone receptors exist as heterodimers *in vivo* and contain between three and seven potential calcium-binding EGF-like motifs. CD97 is predominantly expressed in leukocytes and is markedly upregulated on activated B and T cells (McKnight, A.J. and S. Gordon (1998) J. Leukoc. Biol. 63:271-280).

The third GPCR subfamily is the metabotropic glutamate receptor family. Glutamate is the major excitatory neurotransmitter in the central nervous system. The metabotropic glutamate

receptors modulate the activity of intracellular effectors, and are involved in long-term potentiation (Watson, supra, p.130). The Ca^{2+} -sensing receptor, which senses changes in the extracellular concentration of calcium ions, has a large extracellular domain including clusters of acidic amino acids which may be involved in calcium binding. The metabotropic glutamate receptor family also

5 includes pheromone receptors, the GABA_B receptors, and the taste receptors.

Other subfamilies of GPCRs include two groups of chemoreceptor genes found in the nematodes Caenorhabditis elegans and Caenorhabditis briggsae, which are distantly related to the mammalian olfactory receptor genes. The yeast pheromone receptors STE2 and STE3, involved in the response to mating factors on the cell membrane, have their own seven-transmembrane signature,

10 as do the cAMP receptors from the slime mold Dictyostelium discoideum, which are thought to regulate the aggregation of individual cells and control the expression of numerous developmentally-regulated genes.

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, supra). For instance, retinitis pigmentosa may

15 arise from mutations in the rhodopsin gene. Furthermore, somatic activating mutations in the thyrotropin receptor have been reported to cause hyperfunctioning thyroid adenomas, suggesting that certain GPCRs susceptible to constitutive activation may behave as protooncogenes (Parma, J. et al. (1993) Nature 365:649-651). GPCR receptors for the following ligands also contain mutations associated with human disease: lutenizing hormone (precocious puberty); vasopressin V_2 (X-linked

20 nephrogenic diabetes); glucagon (diabetes and hypertension); calcium (hyperparathyroidism, hypocalcemia, hypercalcemia); parathyroid hormone (short limbed dwarfism); β_3 -adrenoceptor (obesity, non-insulin-dependent diabetes mellitus); growth hormone releasing hormone (dwarfism); and adrenocorticotropin (glucocorticoid deficiency) (Wilson, S. et al. (1998) Br. J. Pharmacol. 125:1387-1392; Stadel, J.M. et al. (1997) Trends Pharmacol. Sci. 18:430-437). GPCRs are also

25 involved in depression, schizophrenia, sleeplessness, hypertension, anxiety, stress, renal failure, and several cardiovascular disorders (Horn, F. and G. Vriend (1998) J. Mol. Med. 76:464-468).

In addition, within the past 20 years several hundred new drugs have been recognized that are directed towards activating or inhibiting GPCRs. The therapeutic targets of these drugs span a wide range of diseases and disorders, including cardiovascular, gastrointestinal, and central nervous system

30 disorders as well as cancer, osteoporosis and endometriosis (Wilson, supra; Stadel, supra). For example, the dopamine agonist L-dopa is used to treat Parkinson's disease, while a dopamine antagonist is used to treat schizophrenia and the early stages of Huntington's disease. Agonists and antagonists of adrenoceptors have been used for the treatment of asthma, high blood pressure, other cardiovascular disorders, and anxiety; muscarinic agonists are used in the treatment of glaucoma and

35 tachycardia; serotonin 5HT_{1D} antagonists are used against migraine; and histamine H₁ antagonists

are used against allergic and anaphylactic reactions, hay fever, itching, and motion sickness (Horn, supra).

Recent research suggests potential future therapeutic uses for GPCRs in the treatment of metabolic disorders including diabetes, obesity, and osteoporosis. For example, mutant V2 vasopressin receptors causing nephrogenic diabetes could be functionally rescued in vitro by co-expression of a C-terminal V2 receptor peptide spanning the region containing the mutations. This result suggests a possible novel strategy for disease treatment (Schöneberg, T. et al. (1996) EMBO J. 15:1283-1291). Mutations in melanocortin-4 receptor (MC4R) are implicated in human weight regulation and obesity. As with the vasopressin V2 receptor mutants, these MC4R mutants are defective in trafficking to the plasma membrane (Ho, G. and R.G. MacKenzie (1999) J. Biol. Chem. 274:35816-35822), and thus might be treated with a similar strategy. The type 1 receptor for parathyroid hormone (PTH) is a GPCR that mediates the PTH-dependent regulation of calcium homeostasis in the bloodstream. Study of PTH/receptor interactions may enable the development of novel PTH receptor ligands for the treatment of osteoporosis (Mannstadt, M. et al. (1999) Am. J. Physiol. 277:F665-F675).

The chemokine receptor group of GPCRs have potential therapeutic utility in inflammation and infectious disease. (For review, see Locati, M. and P.M. Murphy (1999) Annu. Rev. Med. 50:425-440.) Chemokines are small polypeptides that act as intracellular signals in the regulation of leukocyte trafficking, hematopoiesis, and angiogenesis. Targeted disruption of various chemokine receptors in mice indicates that these receptors play roles in pathologic inflammation and in autoimmune disorders such as multiple sclerosis. Chemokine receptors are also exploited by infectious agents, including herpesviruses and the human immunodeficiency virus (HIV-1) to facilitate infection. A truncated version of chemokine receptor CCR5, which acts as a coreceptor for infection of T-cells by HIV-1, results in resistance to AIDS, suggesting that CCR5 antagonists could be useful in preventing the development of AIDS.

The discovery of new G-protein coupled receptors and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, G-protein coupled receptors, referred to collectively as "GCREC" and individually as "GCREC-1" and "GCREC-2." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least

90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-2.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. In one alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:3-4.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an

immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2.

The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), or e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), or e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional GCRC, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group

consisting of SEQ ID NO:1-2, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, or d) an immunogenic fragment of an amino acid sequence selected from the group

consisting of SEQ ID NO:1-2. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional GCREC, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:3-4, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for each polypeptide of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of each polypeptide sequence, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of each polypeptide.

Table 4 lists the cDNA and genomic DNA fragments which were used to assemble each polynucleotide sequence, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for each polynucleotide of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"GCREC" refers to the amino acid sequences of substantially purified GCREC obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of GCREC. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GCREC either by directly interacting with

GCREC or by acting on components of the biological pathway in which GREC participates.

An "allelic variant" is an alternative form of the gene encoding GREC. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding GREC include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as GREC or a polypeptide with at least one functional characteristic of GREC. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GREC, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GREC. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent GREC. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of GREC is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of GREC. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small

molecules, or any other compound or composition which modulates the activity of GCREC either by directly interacting with GCREC or by acting on components of the biological pathway in which GCREC participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind GCREC polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic GCREC, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the

complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding GCRC or fragments of GCRC may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
30	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
35	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu

5	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of GCREC or the polynucleotide encoding GCREC which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:3-4 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:3-4, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:3-4 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:3-4 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:3-4 and the region of SEQ ID NO:3-4 to

which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-2 is encoded by a fragment of SEQ ID NO:3-4. A fragment of SEQ ID NO:1-2 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-2. For example, a fragment of SEQ ID NO:1-2 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-2. The precise length of a fragment of SEQ ID NO:1-2 and the region of SEQ ID NO:1-2 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as

follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

- Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

- Open Gap: 5 and Extension Gap: 2 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

- Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

- Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polymucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized

- 5 after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be
- 10 varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

- Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about
- 15 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY;
- 20 specifically see volume 2, chapter 9.

- High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%.
- 25 Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high
- 30 stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

- The term “hybridization complex” refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one
- 35 nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid

support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

5 "Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

10 An "immunogenic fragment" is a polypeptide or oligopeptide fragment of GCREC which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of GCREC which is useful in any of the antibody production methods disclosed herein or known in the art.

15 The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate. The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of GCREC. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of GCREC.

20 The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

25 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

30 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

35 "Probe" refers to nucleic acid sequences encoding GCREC, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are

isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule.

Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target

- 5 DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,
10 or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold
15 Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

20 Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the
25 PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which
30 sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments,
35 thereby allowing selection of primers that hybridize to either the most conserved or least conserved

regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding GCREC, or fragments thereof, or GCREC itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free,

preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

5 "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a
10 recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term
15 "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic
20 acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The
25 transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989),
30 supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at
35 least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or

greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding

5 polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide

10 polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of

15 the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

20 THE INVENTION

The invention is based on the discovery of new human G-protein coupled receptors (GCREC), the polynucleotides encoding GCREC, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections.

25 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide consensus sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide

30 sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte nucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte

35 polypeptide consensus sequence number (Incyte Polypeptide ID) for each polypeptide of the

invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog.

Table 3 shows various structural features of each of the polypeptides of the invention.

- 5 Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide consensus sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software
- 10 package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

- Table 4 lists the DNA fragments which were used to assemble the full length polynucleotide
- 15 sequences of the invention. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify
- 20 SEQ ID NO:3-4 or that distinguish between SEQ ID NO:3-4 and related polynucleotide sequences. For SEQ ID NO:3-4, column 5 shows Incyte cDNAs (e.g., 5628963H1) and/or GenBank cDNAs or ESTs (e.g., g3333333) which were used to assemble the full length polynucleotide sequences. The cDNA libraries from which the Incyte cDNAs were derived are indicated in parentheses. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In
- 25 some cases, Incyte cDNA coverage redundant with that shown in column 5 was obtained to confirm the final consensus polynucleotide sequence but is not shown. For SEQ ID NO:3, column 5 shows identification numbers (e.g., FL324227_00001) of "stitched" consensus coding sequences which were generated by assembling Incyte cDNA sequences with coding sequences predicted by Genscan analysis of genomic DNA. For SEQ ID NO:3, these "stitched" sequences and/or Genscan-predicted
- 30 coding sequences were used to extend cDNA assemblages to full length. The Genscan-predicted coding sequences may have been edited prior to assembly (see Example IV). Columns 6 and 7 show the nucleotide start (5') and stop (3') positions of the cDNAs and genomic sequences in column 5 relative to their respective full length sequences.

- Table 5 shows the representative cDNA libraries for SEQ ID NO:3-4. The representative
- 35 cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA

sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses GCREC variants. A preferred GCREC variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the GCREC amino acid sequence, and which contains at least one functional or structural characteristic of GCREC.

The invention also encompasses polynucleotides which encode GCREC. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4, which encodes GCREC. The polynucleotide sequences of SEQ ID NO:3-4, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding GCREC. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GCREC. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:3-4. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GCREC.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding GCREC, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring GCREC, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GCREC and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring GCREC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GCREC or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the

peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding GCREC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode GCREC and GCREC derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding GCREC or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:3-4 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding GCREC may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)

Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments
 5 adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.*
 10 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in
 15 length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T)
 20 library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-
 25 specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments
 30 which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GCREC may be cloned in recombinant DNA molecules that direct expression of GCREC, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a
 35 functionally equivalent amino acid sequence may be produced and used to express GCREC.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GCREC-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic

5 oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number
10 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of GCREC, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then
15 subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are
20 optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding GCREC may be synthesized, in whole or in part,
25 using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, GCREC itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the
30 ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of GCREC, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.)
35 The composition of the synthetic peptides may be confirmed by amino acid analysis or by

sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active GCREC, the nucleotide sequences encoding GCREC or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding GCREC. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GCREC. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding GCREC and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GCREC and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding GCREC. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding GCREC. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding GCREC can be achieved using a

multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding GCREC into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for

5 in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of GCREC are needed, e.g. for the production of antibodies, vectors which direct high level expression of GCREC may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

10 Yeast expression systems may be used for production of GCREC. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel,

15 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of GCREC. Transcription of sequences encoding GCREC may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J.

20 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology

25 (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding GCREC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain

30 infective virus which expresses GCREC in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of

35 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are

constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of GCREC in cell lines is preferred. For example, sequences encoding GCREC can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *ap^r* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding GCREC is inserted within a marker gene sequence, transformed cells containing sequences encoding GCREC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GCREC under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding GCREC and that

express GCREC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

5 Immunological methods for detecting and measuring the expression of GCREC using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on GCREC is preferred, but a
10 competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

15 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding GCREC include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GCREC, or any fragments thereof, may be cloned into a vector
20 for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for
25 ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding GCREC may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence
30 and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode GCREC may be designed to contain signal sequences which direct secretion of GCREC through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of
35 the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation,

phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity.

Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the

5 American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GCREC may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric GCREC protein
 10 containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of GCREC activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG,
 15 *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a
 20 proteolytic cleavage site located between the GCREC encoding sequence and the heterologous protein sequence, so that GCREC may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

25 In a further embodiment of the invention, synthesis of radiolabeled GCREC may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

30 Fragments of GCREC may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of GCREC may be synthesized separately and then combined to produce the full length molecule.

35 **THERAPEUTICS**

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GCREC and G-protein coupled receptors. In addition, the expression of GCREC is closely associated with cell proliferation and cancer. Therefore, GCREC appears to play a role in cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections. In the treatment of disorders associated with increased GCREC expression or activity, it is desirable to decrease the expression or activity of GCREC. In the treatment of disorders associated with decreased GCREC expression or activity, it is desirable to increase the expression or activity of GCREC.

Therefore, in one embodiment, GCREC or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a

cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; and a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura,

ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a metabolic disorder such as diabetes, obesity, and osteoporosis; and an infection by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus,

- 5 hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and togavirus.

In another embodiment, a vector capable of expressing GCREC or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those described above.

- 10 In a further embodiment, a pharmaceutical composition comprising a substantially purified GCREC in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those provided above.

- 15 In still another embodiment, an agonist which modulates the activity of GCREC may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GCREC including, but not limited to, those listed above.

- 20 In a further embodiment, an antagonist of GCREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GCREC. Examples of such disorders include, but are not limited to, those cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections described above. In one aspect, an antibody which specifically binds GCREC may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express GCREC.

- 25 In an additional embodiment, a vector expressing the complement of the polynucleotide encoding GCREC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GCREC including, but not limited to, those described above.

- 30 In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

- 35 An antagonist of GCREC may be produced using methods which are generally known in the art. In particular, purified GCREC may be used to produce antibodies or to screen libraries of

pharmaceutical agents to identify those which specifically bind GCREC. Antibodies to GCREC may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GCREC or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GCREC have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of GCREC amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to GCREC may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce GCREC-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte

population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for GCREC may also be generated.

5 For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

10 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between GCREC and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies
15 reactive to two non-interfering GCREC epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for GCREC. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of GCREC-antibody complex
20 divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple GCREC epitopes, represents the average affinity, or avidity, of the antibodies for GCREC. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular GCREC epitope, represents a true measure of affinity. High-affinity antibody
25 preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the GCREC-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of GCREC, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).
30

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml,
35 preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation

of GCREC-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding GCREC, or any
 5 fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding GCREC may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding GCREC. Thus, complementary molecules or fragments may be used to modulate GCREC activity, or to achieve regulation of gene function. Such
 10 technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding GCREC.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted
 15 organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding GCREC. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding GCREC can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding GCREC. Such
 20 constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing
 25 complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding GCREC. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix
 30 pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block
 35 translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze

5 endonucleolytic cleavage of sequences encoding GCREC.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for

10 secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques

15 for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding GCREC. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into

20 cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs

25 and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable

30 for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

35 Any of the therapeutic methods described above may be applied to any subject in need of

such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of GCREC, antibodies to GCREC, and mimetics, agonists, antagonists, or inhibitors of GCREC. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol.

- 5 Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

- 10 Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil,
- 15 or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

- 20 For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

- 25 The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

- 30 After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of GCREC, such labeling would include amount, frequency, and method of administration.

- 35 Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example GCREC or fragments thereof, antibodies of GCREC, and agonists, antagonists or inhibitors of GCREC, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind GCREC may be used for the diagnosis of disorders characterized by expression of GCREC, or in assays to monitor patients being treated with GCREC or agonists, antagonists, or inhibitors of GCREC. Antibodies useful for

diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GCREC include methods which utilize the antibody and a label to detect GCREC in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring GCREC, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GCREC expression. Normal or standard values for GCREC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to GCREC under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of GCREC expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding GCREC may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of GCREC may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GCREC, and to monitor regulation of GCREC levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GCREC or closely related molecules may be used to identify nucleic acid sequences which encode GCREC. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding GCREC, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the GCREC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:3-4 or from genomic sequences including promoters, enhancers, and introns of the GCREC gene.

Means for producing specific hybridization probes for DNAs encoding GCREC include the cloning of polynucleotide sequences encoding GCREC or GCREC derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may

be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

- 5 Polynucleotide sequences encoding GCRC may be used for the diagnosis of disorders associated with expression of GCRC. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including
- 10 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms,
- 15 Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous
- 20 system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial
- 25 nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic
- 30 neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure,
- 35 ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease,

- degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart
- 5 disease, congenital heart disease, and complications of cardiac transplantation; and a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis,
- 10 pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic
- 15 encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; an
- 20 autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes
- 25 mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome,
- 30 systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a metabolic disorder such as diabetes, obesity, and osteoporosis; and an infection by a viral agent classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus,
- 35 hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus,

picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and togavirus. The polynucleotide sequences encoding GCREC may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered GCREC

5 expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding GCREC may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GCREC may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a
10 suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding GCREC in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to
15 monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of GCREC, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding GCREC, under conditions suitable for hybridization or
20 amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

25 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

30 With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development
35 or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding GCREC may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding GCREC, or a fragment of a polynucleotide complementary to the polynucleotide encoding GCREC, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of GCREC include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding GCREC may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the

5 location of the gene encoding GCREC on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as

10 linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other

15 gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among

20 normal, carrier, or affected individuals.

In another embodiment of the invention, GCREC, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes

25 between GCREC and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with GCREC, or fragments thereof,

30 and washed. Bound GCREC is then detected by methods well known in the art. Purified GCREC can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing

35 antibodies capable of binding GCREC specifically compete with a test compound for binding

GCREC. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GCREC.

In additional embodiments, the nucleotide sequences which encode GCREC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs which were used to assemble all or portions of SEQ ID NO:3-4 were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Pharmaceuticals, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A⁺) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the

appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,

- 5 PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

- 10 Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96
- 15 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in

- 20 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.

- 25 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI
- 30 PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the
- 35 cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit

7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMR. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, or Genscan-predicted coding sequences (see Example IV) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences which were subsequently analyzed by querying against databases such as the GenBank protein databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ

ID NO:3-4. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative G-protein coupled receptors were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and Karlin, S. (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and Karlin, S. (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode G-protein coupled receptors, the encoded polypeptides were analyzed by querying against PFAM models for G-protein coupled receptors. Potential G-protein coupled receptors were also identified by homology to Incyte cDNA sequences that had been annotated as G-protein coupled receptors. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. SEQ ID NO:3 were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences to obtain full length polynucleotide sequences using the assembly process described in Example III.

V. Chromosomal Mapping of GCREC Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:3-4 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:3-4 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, or human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-

arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Génethon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses may also be reported as a percentage distribution of libraries in which the transcript encoding GCREC occurred. Analysis involves the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories.

VII. Extension of GCREC Encoding Polynucleotides

The full length polynucleotide sequences of SEQ ID NO:4 were produced by extension of an

appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on

antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the full length polynucleotide sequences of SEQ ID NO:3 are verified using the above procedure.

In like manner, the polynucleotide sequences of SEQ ID NO:3-4 are used to obtain 5' regulatory sequences using the procedure above along with oligonucleotides designed for such extension, and an appropriate genomic library.

VIII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:3-4 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

IX. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

X. Complementary Polynucleotides

Sequences complementary to the GCRC-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GCRC. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of GCRC. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the GCRC-encoding transcript.

XI. Expression of GCRC

Expression and purification of GCRC is achieved using bacterial or virus-based expression systems. For expression of GCRC in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory

element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3).

Antibiotic resistant bacteria express GCREC upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GCREC in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus

- 5 (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding GCREC by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases.
- 10 Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, GCREC is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from GCREC at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified GCREC obtained by these methods can be used directly in the following assay and in the assays shown in Examples XV, XVI, and XVII.

25 XII. Functional Assays

- GCREC function is assessed by expressing the sequences encoding GCREC at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green
- 30
 - 35

Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of GCREC on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GCREC and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding GCREC and other genes of interest can be analyzed by northern analysis or microarray techniques.

20 XIII. Production of GCREC Specific Antibodies

GCREC substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the GCREC amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-GCREC activity by, for example, binding the peptide or GCREC to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring GCREC Using Specific Antibodies

Naturally occurring or recombinant GCREC is substantially purified by immunoaffinity chromatography using antibodies specific for GCREC. An immunoaffinity column is constructed by covalently coupling anti-GCREC antibody to an activated chromatographic resin, such as

- 5 CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing GCREC are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GCREC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt

- 10 antibody/GCREC binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GCREC is collected.

XV. Identification of Molecules Which Interact with GCREC

Molecules which interact with GCREC may include agonists and antagonists, as well as molecules involved in signal transduction, such as G proteins. GCREC, or a fragment thereof, is

- 15 labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) A fragment of GCREC includes, for example, a fragment comprising one or more of the three extracellular loops, the extracellular N-terminal region, or the third intracellular loop. Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GCREC, washed, and any wells with labeled GCREC complex are assayed. Data obtained
- 20 using different concentrations of GCREC are used to calculate values for the number, affinity, and association of GCREC with the candidate ligand molecules.

Alternatively, molecules interacting with GCREC are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

- 25 Potential GCREC agonists or antagonists may be tested for activation or inhibition of GCREC receptor activity using the assays described in sections XVI and XVII. Candidate molecules may be selected from known GPCR agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

Methods for detecting interactions of GCREC with intracellular signal transduction

- 30 molecules such as G proteins are based on the premise that internal segments or cytoplasmic domains from an orphan G protein-coupled seven transmembrane receptor may be exchanged with the analogous domains of a known G protein-coupled seven transmembrane receptor and used to identify the G-proteins and downstream signaling pathways activated by the orphan receptor domains (Kobilka, B.K. et al. (1988) *Science* 240:1310-1316). In an analogous fashion, domains of the orphan
- 35 receptor may be cloned as a portion of a fusion protein and used in binding assays to demonstrate

interactions with specific G proteins. Studies have shown that the third intracellular loop of G protein-coupled seven transmembrane receptors is important for G protein interaction and signal transduction (Conklin, B.R. et al. (1993) Cell 73:631-641). For example, the DNA fragment corresponding to the third intracellular loop of GCREC may be amplified by the polymerase chain reaction (PCR) and subcloned into a fusion vector such as pGEX (Pharmacia Biotech). The construct is transformed into an appropriate bacterial host, induced, and the fusion protein is purified from the cell lysate by glutathione-Sepharose 4B (Pharmacia Biotech) affinity chromatography.

For in vitro binding assays, cell extracts containing G proteins are prepared by extraction with 50 mM Tris, pH 7.8, 1 mM EGTA, 5 mM MgCl₂, 20 mM CHAPS, 20% glycerol, 10 µg of both aprotinin and leupeptin, and 20 µl of 50 mM phenylmethylsulfonyl fluoride. The lysate is incubated on ice for 45 min with constant stirring, centrifuged at 23,000 g for 15 min at 4°C, and the supernatant is collected. 750 µg of cell extract is incubated with GST fusion protein beads for 2 h at 4°C. The GST beads are washed five times with phosphate-buffered saline. Bound G subunits are detected by [³²P]ADP-ribosylation with pertussis or cholera toxins. The reactions are terminated by the addition of SDS sample buffer (4.6% (w/v) SDS, 10% (v/v) β-mercaptoethanol, 20% (w/v) glycerol, 95.2 mM Tris-HCl, pH 6.8, 0.01% (w/v) bromphenol blue). The [³²P]ADP-labeled proteins are separated on 10% SDS-PAGE gels, and autoradiographed. These gels are transferred to nitrocellulose paper, blocked with blotto (5% nonfat dried milk, 50 mM Tris-HCl (pH 8.0), 2 mM CaCl₂, 80 mM NaCl, 0.02% NaN₃, and 0.2% Nonidet P-40) for 1 hour at room temperature, followed by incubation for 1.5 hours with Gα subtype selective antibodies (1:500; Calbiochem-Novabiochem). After three washes, blots are incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit immunoglobulin (1:2000, Cappel, Westchester PA) and visualized by the chemiluminescence-based ECL method (Amersham Corp.).

XVI. Demonstration of GCREC Activity

An assay for GCREC activity measures the expression of GCREC on the cell surface. cDNA encoding GCREC is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using GCREC-specific antibodies, and immunoprecipitated samples are analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of GCREC expressed on the cell surface.

In the alternative, an assay for GCREC activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the rate of DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding GCREC is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The

transiently transfected cells are then incubated in the presence of [³H]thymidine, a radioactive DNA precursor molecule. Varying amounts of GCREC ligand are then added to the cultured cells. Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold GCREC ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of GCREC producing a 50% response level, where 100% represents maximal incorporation of [³H]thymidine into acid-precipitable DNA (McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY, p. 73.)

In a further alternative, the assay for GCREC activity is based upon the ability of GPCR family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) *J. Biol. Chem.* 273:4990-4996). A plasmid encoding full length GCREC is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the attached cells are gently washed with PBS. The cells are then incubated in culture medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand compared to those without ligand are proportional to the amount of GCREC present in the transfected cells.

To measure changes in inositol phosphate levels, the cells are grown in 24-well plates containing 1×10^5 cells/well and incubated with inositol-free media and [³H]myoinositol, 2 μ Ci/well, for 48 hr. The culture medium is removed, and the cells washed with buffer containing 10 mM LiCl followed by addition of ligand. The reaction is stopped by addition of perchloric acid. Inositol phosphates are extracted and separated on Dowex AG1-X8 (Bio-Rad) anion exchange resin, and the total labeled inositol phosphates counted by liquid scintillation. Changes in the levels of labeled inositol phosphate from cells exposed to ligand compared to those without ligand are proportional to the amount of GCREC present in the transfected cells.

XVII. Identification of GCREC ligands

GCREC is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293 which have a good history of GPCR expression and which contain a wide range of G-proteins allowing for functional coupling of the expressed GCREC to downstream effectors. The transformed cells are assayed for activation of the expressed receptors in

the presence of candidate ligands. Activity is measured by changes in intracellular second messengers, such as cyclic AMP or Ca^{2+} . These may be measured directly using standard methods well known in the art, or by the use of reporter gene assays in which a luminescent protein (e.g. firefly luciferase or green fluorescent protein) is under the transcriptional control of a promoter responsive to the stimulation of protein kinase C by the activated receptor (Milligan, G. et al. (1996) Trends Pharmacol. Sci. 17:235-237). Assay technologies are available for both of these second messenger systems to allow high throughput readout in multi-well plate format, such as the adenylyl cyclase activation FlashPlate Assay (NEN Life Sciences Products), or fluorescent Ca^{2+} indicators such as Fluo-4 AM (Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In cases where the physiologically relevant second messenger pathway is not known, GCREC may be coexpressed with the G-proteins $G_{\alpha 15/16}$ which have been demonstrated to couple to a wide range of G-proteins (Offermans, S. and M.I. Simon (1995) J. Biol. Chem. 270:15175-15180), in order to funnel the signal transduction of the GCREC through a pathway involving phospholipase C and Ca^{2+} mobilization. Alternatively, GCREC may be expressed in engineered yeast systems which lack endogenous GPCRs, thus providing the advantage of a null background for GCREC activation screening. These yeast systems substitute a human GPCR and Ga protein for the corresponding components of the endogenous yeast pheromone receptor pathway. Downstream signaling pathways are also modified so that the normal yeast response to the signal is converted to positive growth on selective media or to reporter gene expression (Broach, J.R. and J. Thorner (1996) Nature 384 (supp.):14-16). The receptors are screened against putative ligands including known GPCR ligands and other naturally occurring bioactive molecules. Biological extracts from tissues, biological fluids and cell supernatants are also screened.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-2,
 - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-2,
 - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and
 - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-2.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-2.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:3-4.
- 20 5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
6. A cell transformed with a recombinant polynucleotide of claim 5.
- 25 7. A transgenic organism comprising a recombinant polynucleotide of claim 5.
8. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide
 - 30 comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.
9. An isolated antibody which specifically binds to a polypeptide of claim 1.

35

10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,
- b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a
- 5 polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4,
- c) a polynucleotide sequence complementary to a),
- d) a polynucleotide sequence complementary to b), and
- e) an RNA equivalent of a)-d).

10 11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.

12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:

- 15 a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if
- 20 present, the amount thereof.

13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.

14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.

25 15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

16. A method for treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.

17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 35 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

5

19. A method for treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

10

20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

15

21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

20

22. A method for treating a disease or condition associated with overexpression of functional GCREC, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

25

23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

ABSTRACT OF THE DISCLOSURE

The invention provides human G-protein coupled receptors (GCREC) and polynucleotides which identify and encode GCREC. The invention also provides expression vectors, host cells,

antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of GCREC.

[illegible]

CONFIDENTIAL

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
5628963	1	5628963CD1	3	5628963CB1
1453124	2	1453124CD1	4	1453124CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	5628963CD1	g5353887	2.2e-54	Cysteinyl leukotriene LTD4 receptor [Homo sapiens]
2	1453124CD1	g5525078	1.0e-134	Seven transmembrane receptor [Rattus norvegicus]

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	5628963CD1	346	T71 T344 S13 S15 S176 S186 S151 S241 T265 T273	N20 N26 N30 N181	Signal peptide: M1-G55 Transmembrane domains: R38-Y61, L198-Y221 7 transmembrane receptor (rhodopsin family): G55-Y305 G-protein coupled receptors signature: S115-I166 G-protein coupled receptor: BL00237A: W104-P143 BL00237B: C210-Y221 BL00237C: V240-L266 BL00237D: N297-K313 Rhodopsin-like GPCR superfamily: PR00237A: F40-L64 PR00237B: V73-R94 PR00237C: L118-M140 PR00237D: S154-D175 PR00237E: N202-I225 PR00237F: A245-V269 PR00237G: L287-K313 Chemokine receptor signature: PR00657C: F87-Y97 PR00657D: S115-F128 PR00657G: F307-R321 G-protein coupled receptor: DM00013 P32249 25-324: R38-R321 G-protein coupled receptor: DM00013 P41231 27-322: E34-E310 G-protein coupled receptor: PD000009: K68-I171	SPScan HMMER HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS BLAST-DOMO BLAST-DOMO BLAST-PRODOR

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	1453124CD1	910	T13 T30 T55 T104 S115 S219 S225 S266 S302 S369 T370 S376 S474 S514 S523 T773 S779 T810 T55 T155 S248 T284 S401 S425 T622 T726 S858 T862 S871	N139 N168 N205 N282 N310 N317 N329 N354 N368 N389 N410 N423 N437 N455 N512 N528 N553 N736 N739 N866	Signal_peptide: M1-G19 Signal peptide: M1-G17 Transmembrane domains: L593-W611, M627-D648, L671-I687, L744-L764, K791-V812, V821-L840 Latrophilin/CL-1-like GPS domain: S530-T580 7 transmembrane receptor (secretin family): F582-L845 G-protein coupled receptor: BL00649B: F667-I712 BL00649C: C657-L682 BL00649D: A746-W770 BL00649E: L745-V774 BL00649F: I786-G807 BL00649G: A828-L853 cAMP-type GPCR signature: PR00247A: T589-W611 PR00247C: V661-I687 PR00247D: V701-A719 PR00247E: L744-L764 Secretin-like GPCR superfamily: PR00249A: W587-W611 PR00249C: A659-L682 PR00249E: L745-W770 PR00249F: V789-G809 PR00249G: H820-L841 Hormone receptor EMR1 precursor: DM05221 I37225 347-738: C534-Q847 Hormone receptor EMR1 precursor: DM05221 P48960 347-738: C534-Q847	HMMER SPScan HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS BLAST-DOMO BLAST-DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Hormone receptor EMRI precursor: DM05221 A57172 465-886: E526-K878	BLAST-DOMO
					G-protein coupled receptors family 2: DM00378 P34998 5-438: Q613-H888	BLAST-DOMO
					G-protein coupled receptor: PD000752: V585-Q847	BLAST- PRODOM

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5', Position	3', Position
3	5628963CB1	1625	1-44, 437-1625	5628963H1 (PLACFER01)	1	264
				7163068F8 (PLACNOR01)	1024	1625
				FL324227_00001	137	1184
4	1453124CB1	3446		4862394T8 (PROSTUT09)	2777	3429
				955131R6 (KIDNNOT05)	1959	2484
				4862394T9 (PROSTUT09)	2972	3446
				70485837V1	18	619
				955131T6 (KIDNNOT05)	2416	3054
				60137393B1	1480	1871
				70491324V1	449	791
				70483085V1)	736	1355
				7328527H1	1851	2304
				2188518F6 (PROSNOT26)	1	435
				70485673V1	990	1452
				70484098V1	1242	1841

PI-0072 P

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
3	5628963CB1	PLACNOR01
4	1453124CB1	PROSTUT09

Table 6

Library	Vector	Library Description
PLACNOR01	PCDNA2.1	Library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from placental tissue removed from a Caucasian fetus (donor A), who died after 16 weeks' gestation from fetal demise and hydrocephalus, and a Caucasian male fetus (donor B), who died after 18 weeks' gestation from fetal demise. Serology was positive for anti-cytomegalovirus in donor A and negative in donor B. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time) in donor A. Family history included multiple pregnancies and live births, and an abortion in the mother in donor A.
PROSTUT09	pINCY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma. The patient presented with prostatic inflammatory disease. Patient history included lung neoplasm, and benign hypertension. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease and lung cancer.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, <i>Nucl. Acid Res.</i> , 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Attwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> , 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score>GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPSscan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. ^{supra} ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

SEQUENCE LISTING

<110> Policky, Jennifer L.
 Tribouley, Catherine M.
 Tang, Y. Tom
 Baughn, Mariah R.
 Gaul, Rick
 Khan, Farrah A.
 Nguyen, Danniel B.
 Patterson, Chandra

<120> G-PROTEIN COUPLED RECEPTORS

<130> PI-0072 P

<140> To Be Assigned

<141> Herewith

<160> 4

<170> PERL Program

<210> 1

<211> 346

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5628963CD1

<400> 1

Met	Glu	Arg	Lys	Phe	Met	Ser	Leu	Gln	Pro	Ser	Ile	Ser	Val	Ser	1	5	10	15
Glu	Met	Glu	Pro	Asn	Gly	Thr	Phe	Ser	Asn	Asn	Asn	Ser	Arg	Asn	20	25	30	35
Cys	Thr	Ile	Glu	Asn	Phe	Lys	Arg	Glu	Phe	Phe	Pro	Ile	Val	Tyr	40	45	50	55
Leu	Ile	Ile	Phe	Phe	Trp	Gly	Val	Leu	Gly	Asn	Gly	Leu	Ser	Ile	60	65	70	75
Tyr	Val	Phe	Leu	Gln	Pro	Tyr	Lys	Lys	Ser	Thr	Ser	Val	Asn	Val	80	85	90	95
Phe	Met	Leu	Asn	Leu	Ala	Ile	Ser	Asp	Leu	Leu	Phe	Ile	Ser	Thr	100	105	110	115
Leu	Pro	Phe	Arg	Ala	Asp	Tyr	Tyr	Leu	Arg	Gly	Ser	Asn	Trp	Ile	120	125	130	135
Phe	Gly	Asp	Leu	Ala	Cys	Arg	Ile	Met	Ser	Tyr	Ser	Leu	Tyr	Val	140	145	150	155
Asn	Met	Tyr	Ser	Ser	Ile	Tyr	Phe	Leu	Thr	Val	Leu	Ser	Val	Val	160	165	170	175
Arg	Phe	Leu	Ala	Met	Val	His	Pro	Phe	Arg	Leu	Leu	His	Val	Thr	180	185	190	195
Ser	Ile	Arg	Ser	Ala	Trp	Ile	Leu	Cys	Gly	Ile	Ile	Trp	Ile	Leu	200	205	210	

PI-0072 P

Leu	Leu	Pro	Phe	Phe	Thr	Leu	Ser	Ile	Cys	Tyr	Leu	Leu	Ile	Ile
				215					220					225
Arg	Val	Leu	Leu	Lys	Val	Glu	Val	Pro	Glu	Ser	Gly	Leu	Arg	Val
				230					235					240
Ser	His	Arg	Lys	Ala	Leu	Thr	Thr	Ile	Ile	Ile	Thr	Leu	Ile	Ile
				245					250					255
Phe	Phe	Leu	Cys	Phe	Leu	Pro	Tyr	His	Thr	Leu	Arg	Thr	Val	His
				260					265					270
Leu	Thr	Thr	Trp	Lys	Val	Gly	Leu	Cys	Lys	Asp	Arg	Leu	His	Lys
				275					280					285
Ala	Leu	Val	Ile	Thr	Leu	Ala	Leu	Ala	Ala	Ala	Asn	Ala	Cys	Phe
				290					295					300
Asn	Pro	Leu	Leu	Tyr	Tyr	Phe	Ala	Gly	Glu	Asn	Phe	Lys	Asp	Arg
				305					310					315
Leu	Lys	Ser	Ala	Leu	Arg	Lys	Gly	His	Pro	Gln	Lys	Ala	Lys	Thr
				320					325					330
Lys	Cys	Val	Phe	Pro	Val	Ser	Val	Trp	Leu	Arg	Lys	Glu	Thr	Arg
				335					340					345
Val														

<210> 2

<211> 910

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1453124CD1

<400> 2

Met	Lys	Val	Gly	Val	Leu	Trp	Leu	Val	Ser	Phe	Phe	Thr	Phe	Thr
1				5					10					15
Asp	Gly	His	Gly	Gly	Phe	Leu	Gly	Lys	Asn	Asp	Gly	Ile	Lys	Thr
				20					25					30
Lys	Lys	Glu	Leu	Ile	Val	Asn	Lys	Lys	Lys	His	Leu	Gly	Pro	Val
				35					40					45
Glu	Glu	Tyr	Gln	Leu	Leu	Leu	Gln	Val	Thr	Tyr	Arg	Asp	Ser	Lys
				50					55					60
Glu	Lys	Arg	Asp	Leu	Arg	Asn	Phe	Leu	Lys	Leu	Leu	Lys	Pro	Pro
				65					70					75
Leu	Leu	Trp	Ser	His	Gly	Leu	Ile	Arg	Ile	Ile	Arg	Ala	Lys	Ala
				80					85					90
Thr	Thr	Asp	Cys	Asn	Ser	Leu	Asn	Gly	Val	Leu	Gln	Cys	Thr	Cys
				95					100					105
Glu	Asp	Ser	Tyr	Thr	Trp	Phe	Pro	Pro	Ser	Cys	Leu	Asp	Pro	Gln
				110					115					120
Asn	Cys	Tyr	Leu	His	Thr	Ala	Gly	Ala	Leu	Pro	Ser	Cys	Glu	Cys
				125					130					135
His	Leu	Asn	Asn	Leu	Ser	Gln	Ser	Val	Asn	Phe	Cys	Glu	Arg	Thr
				140					145					150
Lys	Ile	Trp	Gly	Thr	Phe	Lys	Ile	Asn	Glu	Arg	Phe	Thr	Asn	Asp
				155					160					165
Leu	Leu	Asn	Ser	Ser	Ser	Ala	Ile	Tyr	Ser	Lys	Tyr	Ala	Asn	Gly
				170					175					180
Ile	Glu	Ile	Gln	Leu	Lys	Lys	Ala	Tyr	Glu	Arg	Ile	Lys	Gly	Phe
				185					190					195
Glu	Ser	Val	Gln	Val	Thr	Gln	Phe	Arg	Asn	Gly	Ser	Ile	Val	Ala
				200					205					210
Gly	Tyr	Glu	Val	Val	Gly	Ser	Ser	Ser	Ala	Ser	Glu	Leu	Leu	Ser
				215					220					225

3

PI-0072 P

Leu	Met	Leu	Gly	Ile	Leu	Leu	Ala	Tyr	Arg	Ile	Ile	Leu	Val	Phe	
				680					685					690	
His	His	Met	Ala	Gln	His	Leu	Met	Met	Ala	Val	Gly	Phe	Cys	Leu	
				695					700					705	
Gly	Tyr	Gly	Cys	Pro	Leu	Ile	Ile	Ser	Val	Ile	Thr	Ile	Ala	Val	
				710					715					720	
Thr	Gln	Pro	Ser	Asn	Thr	Tyr	Lys	Arg	Lys	Asp	Val	Cys	Trp	Leu	
				725					730					735	
Asn	Trp	Ser	Asn	Gly	Ser	Lys	Pro	Leu	Leu	Ala	Phe	Val	Val	Pro	
				740					745					750	
Ala	Leu	Ala	Ile	Val	Ala	Val	Asn	Phe	Val	Val	Val	Leu	Leu	Val	
				755					760					765	
Leu	Thr	Lys	Leu	Trp	Arg	Pro	Thr	Val	Gly	Glu	Arg	Leu	Ser	Arg	
				770					775					780	
Asp	Asp	Lys	Ala	Thr	Ile	Ile	Arg	Val	Gly	Lys	Ser	Leu	Leu	Ile	
				785					790					795	
Leu	Thr	Pro	Leu	Leu	Gly	Leu	Thr	Trp	Gly	Phe	Gly	Ile	Gly	Thr	
				800					805					810	
Ile	Val	Asp	Ser	Gln	Asn	Leu	Ala	Trp	His	Val	Ile	Phe	Ala	Leu	
				815					820					825	
Leu	Asn	Ala	Phe	Gln	Gly	Phe	Phe	Ile	Leu	Cys	Phe	Gly	Ile	Leu	
				830					835					840	
Leu	Asp	Ser	Lys	Leu	Arg	Gln	Leu	Leu	Phe	Asn	Lys	Leu	Ser	Ala	
				845					850					855	
Leu	Ser	Ser	Trp	Lys	Gln	Thr	Glu	Lys	Gln	Asn	Ser	Ser	Asp	Leu	
				860					865					870	
Ser	Ala	Lys	Pro	Lys	Phe	Ser	Lys	Pro	Phe	Asn	Pro	Leu	Gln	Asn	
				875					880					885	
Lys	Gly	His	Tyr	Ala	Phe	Ser	His	Thr	Gly	Asp	Ser	Ser	Asp	Asn	
				890					895					900	
Ile	Met	Leu	Thr	Gln	Phe	Val	Ser	Asn	Glu						
				905					910						

<210> 3

<211> 1625

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5628963CB1

<400> 3

ggcagcagaa	gccagggcag	ctgaaagaca	gagaccttca	gggaaaacta	ggttccaaga	60
tggctgaata	ggaagagctc	cagtcctgcag	atcccagtg	gagcaacgtg	gaagacgggt	120
gatttctgca	tttccaactg	agcatggaga	gaaaatttat	gtccttgcaa	ccatccatct	180
ccgtatcaga	aatggaacca	aatggcacct	tcagcaataa	caacagcagg	aactgcacaa	240
ttgaaaactt	caagagagaa	tttttcccaa	ttgtatatct	gataatattt	ttctggggag	300
tcttgggaaa	tgggttgctc	atatatgttt	tcctgcagcc	ttataagaag	tccacatctg	360
tgaacgtttt	catgctaaat	ctggccattt	cagatctcct	gttcataagc	acgcttccct	420
tcagggctga	ctattatctt	agaggctcca	attggatatt	tggagacctg	gcctgcagga	480
ttatgtctta	ttccttgat	gtcaacatgt	acagcagtat	ttatttcttg	accgtgctga	540
gtgttgctgc	tttctggca	atgggtcacc	cctttcggct	tctgcatgtc	accagcatca	600
ggagtgcctg	gacccctctg	gggatcatat	ggatccttat	catggcttcc	tcaataatgc	660
tcctggacag	tggctctgag	cagaacggca	gtgtcacatc	atgcttagag	ctgaatctct	720
ataaaattgc	taagctgcag	accatgaact	atattgcctt	ggtggtgggc	tgcttgcctg	780
cattttttcac	actcagcatc	tggtatctgc	tgatcattcg	ggttctgtta	aaagtggagg	840
tcccagaatc	ggggtgcgg	gtttctcaca	ggaaggcact	gaccaccatc	atcatcacct	900
tgatcatctt	cttcttggtg	ttcctgcctt	atcacacact	gaggaccgtc	cacttgacga	960
catggaaagt	gggtttatgc	aaagacagac	tgcataaagc	tttggttatc	acactggcct	1020
tggcagcagc	caatgcctgc	ttcaatcttc	tgctctatta	ctttgctggg	gagaatttta	1080

```
aggacagact aaagtctgca ctcagaaaag gccatccaca gaaggcaaag acaaagtgtg 1140
ttttccctgt tagtggtgtg ttgagaaagg aaacaagagt ataaggagct cttagatgag 1200
acctgttctt gtatccttgt gtccatcttc attcactcat agtctccaaa tgactttgta 1260
tgttacatca ctcccaacaa atgttgattc ttaatattta gttgaccatt acttttgta 1320
ataagaccta cttcaaaaat tttattcagt gtattttcag ttgttgagtc ttaatgaggg 1380
atacaggagg aaaaatccct actagagtcc tgtgggttga aatatcagac tgggaaaaaa 1440
tgcaaagcac attggatcct acttttcttc agatattgaa ccagatctct ggcccatcag 1500
gctttctaaa ttcttcaaaa gagccacaac ttccccagct tctccagctc cctgtctctc 1560
ttcaatccct tgagatatag ccaactaacg acgctactgg aagcctttcc agcacaatgg 1620
cggcc
```

<210> 4

<211> 3446

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1453124CB1

<400> 4

```
cagtgaagcct gtgttcatgc cagtgaagctg ctgtggctca gatactgata ctttctttcc 60
aaacagcata agaagtgatt gagccacaag tatactgaag gaagggctcc ctcgagttgt 120
ggtgtgaaga gataaatcac cagtcacaga ctatgcaccc gactgctgct gttcagttca 180
ggggaaatga aagttggagt gctgtggctc gtttctttct tcaccttcac tgacggccac 240
ggtggcttcc tggggaaaaa tgatggcatc aaaacaaaaa aagaactcat tgtgaataag 300
aaaaaacatc taggccagat cgaagaatat cagctgctgc ttcaggtgac ctatagagac 360
tccaaggaga aaagagactt gagaaatttt ctgaagctct tgaagcctcc attattatgg 420
tcacatgggc taattagaat tatcagagca aaggctacca cagactgcaa cagcctgaat 480
ggagtccctgc agtgtacctg tgaagacagc tacacctggg ttctccctc atgccttgat 540
cccagaact gctaccttca cacggctgga gcactcccaa gctgtgaatg tcactctaac 600
aacctcagcc agagtgtcaa tttctgtgag agaacaaga tttggggcac tttcaaaatt 660
aatgaaaggt ttacaaaatga ccttttgaat tcactctctg ctatatactc caaatatgca 720
aatggaattg aaattcaact taaaaaagca tatgaaagaa ttaaagggtt tgagtcgggt 780
caggtcaccc aatttcgaaa tggagcatc gttgctgggt atgaagttgt tggctccagc 840
agtgcactctg aactgctgtc agccattgaa catgttgccg agaaggctaa gacagccctt 900
cacaagctgt ttccattaga agacggctct ttcagagtgt tcggaaaagc ccagtgtaat 960
gacattgtct ttggatttgg gtccaaggat gatgaatata ccctgccttg cagcagtggtc 1020
tacaggggaa acatcacagc caagtgtgag tcctctgggt ggcaggtcat caggagagact 1080
tgtgtgctct ctctgcttga agaactgaac aagaatttca gtatgattgt aggcaatgcc 1140
actgaggcag ctgtgtcatc cttcgtgcaa aatctttctg tcactattcg gcaaaaccca 1200
tcaaccacag tggggaatct ggcttcgggt gtgtcgatcc tgagcaatat ttcactctctg 1260
tcaactggcca gccatttcag ggtgtccaat tcaacaatgg aggatgtcat cagtatatgt 1320
gacaatatcc ttaattcagc ctcaagtaac aactggacag tcttactgcy ggaagaaaag 1380
tatgccagct cacggttact agagacatta gaaaacatca gcactctggt gcctccgaca 1440
gctcttctct tgaatttttc tcggaaattc attgactgga aagggtattcc agtgaacaaa 1500
agccaactca agaggggtta cagctatcag attaaaatgt gtcccaaaa tacatctatt 1560
cccatcagag gccgtgtgtt aattgggtca gaccaattcc agagatccct tccagaaact 1620
attatcagca tggcctcgtt gactctgggg aacattctac ccgtttccaa aaatggaat 1680
gctcaggtca atggacctgt gatatccacg gttattcaaa actattccat aaatgaagtt 1740
ttctattttt tttccaagat agagtcaaac ctgagccagc ctcatgtgtg gttttgggat 1800
ttcagtcatt tgcagtggaa cgatgcaggc tgccacctag tgaatgaaac tcaagacatc 1860
gtgacgtgcc aatgtactca cttgacctcc ttctccatat tgatgtcacc ttttgtcccc 1920
tctacaatct tccccgttgt aaaatggatc acctatgtgg gactgggtat ctccattgga 1980
agtctcattt atgcctgat catcgaggct ttgttttggg agcagattaa aaaaagccaa 2040
acctctcaca cacgtcgtat ttgcatgggt aacatagccc tgtccctctt gattgctgat 2100
gtctggttta ttgttggtgc cacagtggac accacggtga acccttcttg agtctgcaca 2160
gctgctgtgt tctttacaca cttcttctac ctctctttgt tctctggat gctcatgctt 2220
ggcatcctgc tggcttaccg gatcatcctc gtgttccatc acatggccca gcatttgatg 2280
atggctgttg gattttgcct gggttatggg tgccctctca ttatatctgt cattaccatt 2340
```

```

gctgtcacgc aacctagcaa tacctacaaa aggaaagatg tgtgttggct taactgggcc 2400
aatggaagca aaccactcct ggcttttgtt gtccctgcac tggctattgt ggctgtgaac 2460
ttcgttgtgg tgctgctagt tctcacaaaag ctctggaggc cgactgttgg ggaaagactg 2520
agtcgggatg acaaggccac catcatccgc gtggggaaga gcctcctcat tctgacccct 2580
ctgctagggc tcacctgggg ctttgggaata ggaacaatag tggacagcca gaatctggct 2640
tggcatgtta tttttgcttt actcaatgca ttccagggat tttttatctt atgctttgga 2700
atactcttgg acagtaagct gcgacaactt ctgttcaaca agttgtctgc ctttaagtct 2760
tggaagcaaa cagaaaagca aaactcatca gatttatctg ccaaaccctaa attctcaaag 2820
cctttcaacc cactgcaaaa caaaggccat tatgcatttt ctcatactgg agattcctcc 2880
gacaacatca tgctaactca gtttgtctca aatgaataag gcaaggaatc ataaaatcaa 2940
gaaaaaattt ccagaacaac ttgacattta gagacaaatg tcaatgaaga aattatgctc 3000
agtattcgat cgggttttct gatttagggg tctgggaata aaacaagaat gtctcagtgg 3060
cttcattact gctccctttt gtcttcaatt aaatgaaaag aagatttatt tccatgtgat 3120
ttgattcaaa gaaagtgtc cataaatgca gaagagtagg ttttgttggg aatcgtgtca 3180
gttgaccctt gaccataaaa tatggtttct attttcataa aacagcatta ttcacatggc 3240
atttccaata atctggattg aaggaagaaa attttatgaa atagctttag ataaattaat 3300
aggccacgtt cattttcttg tcaaaaagtt actggtgggg ggatgggtggg aaaaagttaa 3360
tagtgcaaat ttcttagaga aaaaaccatt tctctttcaa attttccagt tgaattttat 3420
gttcgctttt gcttcttagg ttcata 3446

```

CCOE113"10300T03

Characterization of the Human Cysteinyl Leukotriene 2 Receptor*

Received for publication, April 25, 2000, and in revised form, June 5, 2000
Published, JBC Papers in Press, June 12, 2000, DOI 10.1074/jbc.M003490200

Christopher E. Heise†§, Brian F. O'Dowd¶, David J. Figueroa¶, Nicole Sawyer**, Tuan Nguyen¶, Dong-Soon Im‡, Rino Stocco**, Julie N. Bellefeuille**, Mark Abramovitz**, Regina Cheng¶, David L. Williams Jr.¶, Zhizhen Zeng¶, Qingyun Liu¶, Lei Ma¶, Michelle K. Clements¶, Nathalie Coulombe**, Yuan Liu‡‡, Christopher P. Austin¶, Susan R. George¶, Gary P. O'Neill**, Kathleen M. Metters**, Kevin R. Lynch‡§§, and Jilly F. Evans¶¶¶

From the †Department of Pharmacology, University of Virginia, School of Medicine, Charlottesville, Virginia 22908, the ¶Department of Pharmacology, University of Toronto, Medical Sciences Building, Toronto, Ontario M5S 1A8, Canada, Centre for Addiction and Mental Health, Toronto, Ontario M5S 2S1, Canada, the Departments of ¶Pharmacology and ‡‡Bioinformatics, Merck & Co., Inc., West Point, Pennsylvania 19486, and the **Department of Biochemistry and Molecular Biology, Merck Frosst Canada & Co., Pointe-Claire-Dorval, Quebec H9R 4P8, Canada

The contractile and inflammatory actions of the cysteinyl leukotrienes (CysLTs), LTC₄, LTD₄, and LTE₄, are thought to be mediated through at least two distinct but related CysLT G protein-coupled receptors. The human CysLT₁ receptor has been recently cloned and characterized. We describe here the cloning and characterization of the second cysteinyl leukotriene receptor, CysLT₂, a 346-amino acid protein with 38% amino acid identity to the CysLT₁ receptor. The recombinant human CysLT₂ receptor was expressed in *Xenopus* oocytes and HEK293T cells and shown to couple to elevation of intracellular calcium when activated by LTC₄, LTD₄, or LTE₄. Analyses of radiolabeled LTD₄ binding to the recombinant CysLT₂ receptor demonstrated high affinity binding and a rank order of potency for competition of LTC₄ = LTD₄ >> LTE₄. In contrast to the dual CysLT₁/CysLT₂ antagonist, BAY u9773, the CysLT₁ receptor-selective antagonists MK-571, montelukast (SingulairTM), zafirlukast (AccolateTM), and pranlukast (OnonTM) exhibited low potency in competition for LTD₄ binding and as antagonists of CysLT₂ receptor signaling. CysLT₂ receptor mRNA was detected in lung macrophages and airway smooth muscle, cardiac Purkinje cells, adrenal medulla cells, peripheral blood leukocytes, and brain, and the receptor gene was mapped to chromosome 13q14, a region linked to atopic asthma.

The cysteinyl leukotrienes (CysLTs),¹ LTC₄, LTD₄, and LTE₄, previously known as slow reacting substance of anaphylaxis, or SRS-A, are derived from arachidonic acid via oxygenation and dehydration by 5-lipoxygenase followed by specific glutathione addition by LTC₄ synthase (1). The CysLTs medi-

ate their biological actions through two pharmacologically defined G-protein-coupled receptors (GPCRs), named the CysLT₁ and CysLT₂ (2, 3). The recent cloning and characterization of the human CysLT₁ receptor confirmed the previous pharmacological data (Refs. 4 and 5; GenBankTM accession nos. AF 119711 and AF 133266). LTD₄ is the preferred endogenous ligand for the CysLT₁ receptor, and activation of the receptor results in an elevation of intracellular calcium (4, 5). The gene for the CysLT₁ receptor has been mapped to human chromosome Xq13-q21 (4, 5). The CysLT₁ receptor is the molecular target of the anti-asthmatic drugs montelukast (SingulairTM), zafirlukast (AccolateTM), and pranlukast (OnonTM) that have both anti-bronchoconstrictive and anti-inflammatory actions (6–8). All known CysLT receptor antagonists, except BAY u9773 (a nonselective antagonist at CysLT₁ and CysLT₂ receptors) selectively antagonize activation of the CysLT₁ receptor (6–9). CysLT₁ receptor mRNA and protein are expressed on human lung smooth muscle cells and tissue macrophages and on peripheral blood monocytes and eosinophils (4).²

The CysLT₂ receptor has been documented pharmacologically to be expressed in guinea pig trachea and ileum, ferret trachea and spleen, sheep bronchus, and human pulmonary and saphenous vein preparations (2, 10, 11). At the CysLT₂ receptor subtype, the agonist potency rank order is LTC₄ = LTD₄ >> LTE₄, and LTE₄ is a partial agonist (2). We describe here the molecular cloning and characterization of the human CysLT₂ receptor.

EXPERIMENTAL PROCEDURES

Materials—LTD₄ and LTC₄ were from Cayman (Ann Arbor, MI); LTB₄, LTE₄, and BAY u9773 were from BIOMOL (Plymouth Meeting, PA); 1-oleoyl lysophosphatidic acid was from Avanti Polar Lipids (Alabaster, AL). MK-571, montelukast, pranlukast, and zafirlukast were synthesized by the Department of Medicinal Chemistry at Merck Frosst, and [³H]LTD₄ (146 Ci/mmol) was from NEN Life Science Products.

Cloning of HG57, the CysLT₂ Receptor—A partial rat EST (accession no. ai178926) encoding a GPCR fragment, with 40% identity to the human CysLT₁ receptor, was found during a routine search of updates to the GenBankTM data base of expressed sequence tags using the FAST_PAN program (12). The cognate cDNA was retrieved from the American Type Culture Collection and was used to screen a rat brain cDNA library as reported previously (13). Two cDNAs were obtained that had a translational open reading frame of 981 nucleotides. This putative receptor, designated by the expressed sequence tags clone name, RSPBT32, was 40% identical to the human CysLT₁ receptor but 16 and 21 amino acids shorter at the N- and C-terminal ends, respec-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by NIGMS, National Institutes of Health (NIH), National Research Service Award predoctoral traineeship T32 GM07055.

¶¶ The work in this author's laboratory was supported by NIGMS, NIH, Grant R01 GM52722 and a research contract from Merck Research Laboratories.

¶¶¶ To whom correspondence should be addressed: Dept. of Pharmacology, Merck & Co., Inc., 770 Sumneytown Pike, P.O. Box 4, West Point, PA 19486. Tel.: 215-652-1254; Fax: 215-993-4007; E-mail: jilly_evans@merck.com.

¹ The abbreviations used are: CysLT, cysteinyl leukotriene; LTC₄, leukotriene C₄; LTD₄, leukotriene D₄; LTE₄, leukotriene E₄; SSPE, saline/sodium phosphate/EDTA; GPCR, G-protein-coupled receptor.

² D. J. Figueroa, unpublished data.

tively. The RSPBT32 protein sequence was added to the FAST_PAN GPCR query set, and during a subsequent search of GenBank™ releases, we found a human genome survey sequence deposition (accession no. aq001459) encoding a peptide fragment that was about 70% identical to RSPBT32. The genome survey sequence was used to design oligonucleotide primers for the polymerase chain reaction, and ultimately we isolated a human genomic DNA fragment that contained a 1041-nucleotide open reading frame encoding a 346-amino acid protein that we named HG57. The amplified product was subcloned into the TA cloning site of the mammalian cDNA expression vector pCR3.1 (Invitrogen, Carlsbad, CA), and the sequence of the inserted DNA was determined.

Xenopus laevis Oocyte Expression—HG57 (CysLT₂) cRNA was produced by *in vitro* transcription (T7 mMessage mMachine, Ambion) using T7 RNA polymerase in the presence of a capping analogue. 20–40 ng of the capped cRNA was injected into stage V-VI *X. laevis* oocytes as described previously (14, 15). Compounds were delivered as a 30-μl aliquot over a period of 1–2 s, and the recording chamber was washed with 30% methanol between oocyte assays to remove residual ligand. *X. laevis* were purchased from Xenopus I (Ann Arbor, MI).

Aequorin Luminescence Functional Assay—Human embryonic kidney (HEK) 293 cells stably expressing the SV40 large T antigen, designated HEK 293T cells, were transfected with HG57 (CysLT₂)-pCR3.1 or pCR3.1 and AEQ-pCDM plasmids (5 μg of each DNA per 75-cm² culture flask), using the LipofectAMINE™ PLUS reagent (Life Technologies, Inc.) following the manufacturer's instructions. Cells were prepared subsequently for use in the aequorin luminescence functional assay as described previously (16). Briefly, agonists in dimethyl sulfoxide or ethanol were serially diluted, in duplicate, in a white 96-well cliniplate FB (Labsystems) using a Biomek in a final volume of 100 μl in phosphate-buffered saline (with 70 mg of CaCl₂) so that the final solvent concentration was ≤1%. The plate was then loaded into the Luminoskan RS plate-reading luminometer (Labsystems, Needham Heights, MA), and wells were tested sequentially. Cells (~2–5 × 10⁴) in 100 μl of Ham's F12 medium were injected into the well, and light emission was recorded over 30 s (peak 1). The cells were then lysed by injection of 25 μl of 0.9% Triton X-100 solution in H₂O, and light emission was measured for an additional 10 s (peak 2). For the antagonist, the plate was divided into four series. Series 1 contained a leukotriene C₄ or leukotriene D₄ control curve into which transfected cells were injected in the absence of the antagonist, while series 2–4 contained leukotriene C₄ or leukotriene D₄ curves, into which cells that had been preincubated for 15 min in the presence of a given concentration of antagonist were injected.

Fractional luminescence for each well was determined by dividing the area under peak 1 by the total area under peaks 1 and 2 ($P1/(P1 + P2)$). These calculations were performed using the Lakan Controller program, and data files were analyzed using the LDAM software employing a modified version of the Levenberg-Marquardt four-parameter curve-fitting algorithm to calculate EC₅₀ values (14). For the antagonist tests, the EC₅₀ values were then used to generate a K_B value using Schild plot analysis.

Radioligand Binding Assays—COS-7 cell transfection, harvesting, and membrane preparation were done as previously reported (17). Tritiated LTD₄ binding assays were performed as described (17) with the exception that the reaction was initiated by the addition of 250 μg of membrane protein and that 4 mM acivicin replaced serine-borate in equilibrium competition assays with LTC₄. LTD₄-specific binding was calculated by subtracting nonspecific binding, determined in the presence of 1 μM LTD₄, from total binding. Specific binding accounted for 50–60% of the total binding and was linear with respect to the concentrations of radioligand and protein present in the incubation. Total binding represented less than 10% of the radioligand added to the incubation. In these conditions, average total and nonspecific binding of [³H]LTD₄ for 250 μg of HG57 (CysLT₂) receptor membrane protein was 2600 and 1200 dpm, respectively.

Northern Blot Analysis—RNAs from several human tissues were extracted as described previously (18). Total RNA was extracted by the method of Chomczynski and Sacchi (19), and poly(A)⁺ RNA was isolated using oligo(dT)-cellulose spin columns (Amersham Pharmacia Biotech). RNA was denatured and size-fractionated on a 1% formaldehyde-agarose gel, transferred onto nylon membrane, and immobilized by UV irradiation. The blots were hybridized with a ³²P-labeled DNA fragment encoding the CysLT₂ receptor, washed with 2× SSPE and 0.1% SDS at room temperature for 15 min, washed again with 0.5× SSPE and 0.1% SDS at 50 °C for 45 min, subjected to four final washings with 0.2× SSPE and 0.1% SDS at 50 °C each for 1 h, and exposed to x-ray film at –70 °C in the presence of an intensifying screen for 3–7 days.

In Situ Hybridization Analysis and Immunohistochemistry—The oligonucleotide antisense probes used for the *in situ* studies were 5'-C-CAGGCACTCCTGATGCT-3' and 5'-CCCACCACCAAGGAATA-3', and their complementary sequences were used as sense probes. Tailing of oligonucleotides with biotin-16-dUTP (Roche Molecular Biochemicals) was carried out as described by the manufacturer except for substitution of digoxigenin-dUTP with biotin-16-dUTP in the tailing reaction. *In situ* hybridization was carried out on 6-μm cryostat sections of surgical human lung specimens (National Disease Research Interchange, Philadelphia, PA) using 2 pmol/ml labeled oligonucleotides for 18 h at 37 °C. Bound probe was visualized using Texas Red Tyramide Signal Amplification detection reagent (NEN Life Science Products) according to the manufacturer's instructions. Peripheral blood mononuclear cells were isolated from buffy coat preparations by centrifugation over lymphocyte separation medium (ICN). T cells were resorted by incubation of the peripheral blood mononuclear cells with neuraminidase-treated sheep red blood cells and pelleted through lymphocyte separation medium. The sheep red blood cells were removed by lysis with ACK lysis buffer (Life Technologies). T cell-depleted peripheral blood mononuclear cells accumulated at the interface of the lymphocyte separation medium. Eosinophils were prepared from peripheral blood from a donor known to have elevated peripheral blood eosinophils but no history of asthmatic disease. Erythrocytes were removed by hypotonic lysis of the pelleted cells, followed by negative selection with anti-CD16 microbeads (Miltenyi Biotec) according to the manufacturer's instructions. The purity of eosinophil preparations was >90% with some contaminating monocytes and neutrophils. All preparations of cells were resuspended in O.C.T. compound (Miles Scientific) and fresh frozen prior to *in situ* hybridization. *In situ* hybridization on these cell preparations was carried out exactly as described above for tissue sections. Sections and cells were counterstained with 4,6-diamidino-2-phenylindole (Molecular Probes, Inc., Eugene, OR) to visualize cell nuclei, and images were digitally acquired and reassembled using a MicroMax CCD camera (Princeton Instruments) and the Metamorph imaging program (Universal Imaging). Cell nuclei are seen in teal blue pseudocolor or in conventional 4,6-diamidino-2-phenylindole blue. Immunohistochemistry was performed after *in situ* hybridization, with antibodies LN5 (Zymed Laboratories Inc.), CD14 (Serotec), and Protein Gene Product 9.5 (Biomedica Corp.) at the manufacturer's specified dilutions for 2 h at room temperature. Primary antibody markers were detected with either a fluorescein isothiocyanate-conjugated donkey anti-mouse IgG (Jackson ImmunoResearch) or a fluorescein isothiocyanate-conjugated donkey anti-rat IgG (Jackson ImmunoResearch). Colocalization of red (*in situ*) and green (antibody) staining is seen as yellow fluorescence.

Chromosomal Localization—Chromosomal mapping of the CysLT₂ receptor gene was performed using the GeneBridge 4 panel, consisting of 93 radiation hybrid clones (Research Genetics, Huntsville, AL). Two primer pairs were designed from the HG57 sequence to amplify each of two different regions of the gene. Primer pair 1 sequences (5'-AATGG-CACCTTCAGCAATAACA-3' (forward) and 5'-TGGACAACCCATTTTCC-AAGAC-3' (reverse)) produced a band of ~100 base pairs; primer pair 2 sequences (5'-AGACTGCATAAAGCTTTGGTTATC-3' (forward) and 5'-ATACTCTTGTTTCCTTTCTCAACC-3' (reverse)) produced a band of ~180 base pairs. Polymerase chain reaction was performed with AmpliTaq Gold (Perkin-Elmer) using the following cycling parameters: 94 °C for 9 min; 94 °C for 2 min, 62 °C for 30 s, 72 °C for 1 min (32 cycles), and 72 °C for 7 min. Results were submitted to the Whitehead Institute Genome center server (available on the World Wide Web) and confirmed by QUANTUM Somatic Cell Hybrid PCRable Panel (QUANTUM) and Stanford G3 Radiation Hybrid Panel (Research Genetics).

RESULTS AND DISCUSSION

Sequence Comparison—Phylogenetic analysis (Fig. 1A) showed that the human orphan GPCR HG57 (identified here as the CysLT₂ receptor) was 73% identical to the rat orphan GPCR RSBPT32 sequence, 38% identical to the human CysLT₁ receptor sequence, and 33–35% identical to the human orphan GPCRs GPR17 and GPR23. Despite its high similarity to HG57, we have been unable to demonstrate CysLT activation of the rat RSBPT32 receptor (data not shown). The deduced amino acid sequence of HG57 (the CysLT₂ receptor), portrayed as a putative seven-transmembrane domain protein in three-dimensional helical representation, is shown in Fig. 1B (GenBank™ sequence AF254664). The primary sequence of the CysLT₂ receptor is quite dissimilar to another leukotriene re-

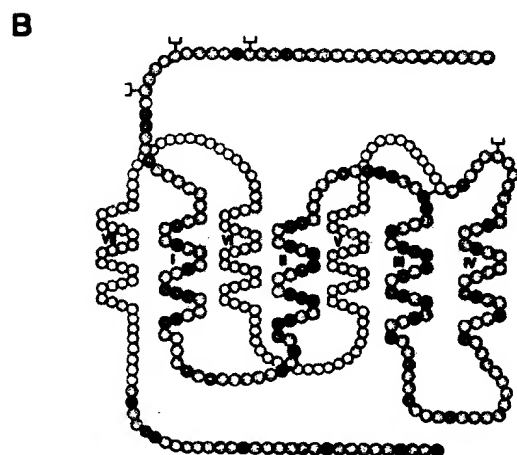
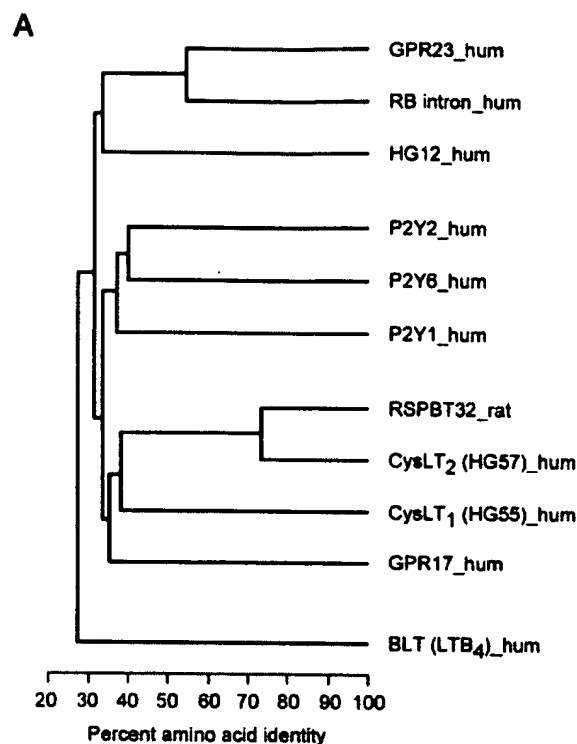


FIG. 1. Amino acid homologies and sequence representation of HG57, the human CysLT₂ receptor. A, phylogenetic tree representing amino acid sequence similarities between the human CysLT₂ (HG57) receptor (accession no. AF254664) and related GPCRs from family 1. These include CysLT₁_hum (AF119711), RSPBT32_rat (AI178926), GPR17_hum (U33447), P2Y1_hum (Z49205), P2Y2_hum (U07225), P2Y6_hum (X97058), HG12_hum (af_000545), RB intron_hum (L11910), GPR23_hum (U66578), and BLT (LTB₄)_hum (U41070). The Wisconsin Sequence Analysis Package (version 10.0) was used in the analysis. Branch lengths along the x axis are inversely proportional to percentage of amino acid identity. B, three-dimensional representation of the amino acid sequence of the CysLT₂ receptor. Glycosylation and phosphorylation sites are indicated, as are the putative seven-transmembrane helices. Amino acids identical in the CysLT₁ and CysLT₂ receptors are indicated by red circles.

ceptor, the BLT high affinity LTB₄ receptor (Ref. 20; Fig. 1A). Although there is no record of HG57 in the expressed sequence tags division of GenBank™, an identical open reading frame is contained as an HTGS deposition (accession no. AL137118).

Xenopus Oocyte Functional Activation—The CysLTs LTD₄ and LTC₄ produced calcium-dependent chloride flux in HG57 (CysLT₂) receptor cRNA-injected *X. laevis* oocytes and were desensitized to subsequent challenge by CysLTs (Fig. 2A and

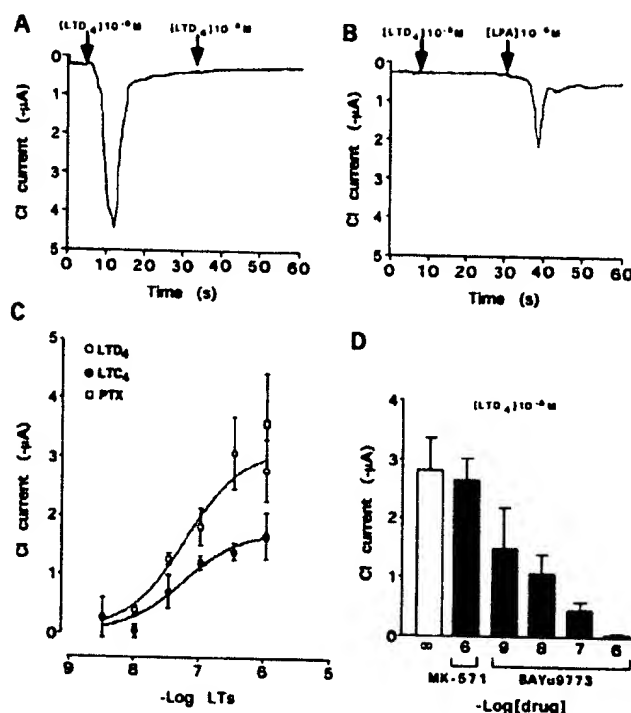


FIG. 2. Functional activation of HG57, the human CysLT₂ receptor expressed in *X. laevis* oocytes. Oocyte calcium-dependent chloride conductance is shown in response to HG57 (CysLT₂) receptor cRNA (A, C, and D) or saline (B) and challenge with LTD₄, LTC₄, or lysophosphatidic acid. C, responses elicited after challenge with LTD₄ (○) or LTC₄ (●) ($n = 3-8$ oocytes per point \pm S.E.), and HG57 (CysLT₂) receptor-injected oocytes pretreated with pertussis toxin (□; 100 ng/ml, 24 h, $n = 3-8$ oocytes). D, responses from HG57 (CysLT₂) receptor cRNA-injected oocytes ($n = 3-8$) to 1 μ M LTD₄ pretreated with MK-571 or BAY u9773.

data not shown). Control oocytes injected with saline or other GPCR cRNAs, including those encoding RSPBT32 and GPR17, showed no response to LTD₄ or LTC₄ but were able to respond to lysophosphatidic acid challenge through an endogenous lysophosphatidic acid receptor (Fig. 2B). LTD₄ and LTC₄ produced dose-dependent activation of calcium-dependent chloride flux in HG57 (CysLT₂) receptor cRNA-injected *Xenopus* oocytes, and the maximal response was not significantly reduced by preincubation with pertussis toxin (Fig. 2C), indicative that CysLT₂ did not couple to G_i α in this system. The LTD₄ or LTC₄ activation of calcium-dependent chloride flux in the HG57 (CysLT₂)-expressing oocytes was not blocked by the CysLT₁ receptor-selective antagonist, MK-571, but was blocked by the dual CysLT₁/CysLT₂ receptor antagonist BAY u9773 (Fig. 2D and data not shown).

Transient Expression and Functional Activation in Mammalian Cells—In HG57 (CysLT₂) and aequorin-co-expressing HEK293T cells, LTD₄ and LTC₄ were equipotent agonists, while LTE₄ behaved as a partial agonist, and LTB₄ was inactive up to 3 μ M (Fig. 3A). These agonists show no response in vector-transfected aequorin-expressing HEK293T cells (data not shown). BAY u9773 acted as a noncompetitive antagonist of both LTC₄- and LTD₄-challenged CysLT₂ receptor-expressing HEK293T cells (Fig. 3B). The selective CysLT₁ receptor antagonists, MK-571, montelukast, zafirlukast, and pranlukast, showed no significant antagonism of the CysLT₂ receptor up to 1 μ M concentration (Fig. 3C).

Radioligand Binding Characterization—Saturation analysis of [³H]LTD₄-specific binding to COS-7 cell membranes transiently expressing HG57 (CysLT₂) was performed in two separate experiments with two different membrane preparations.

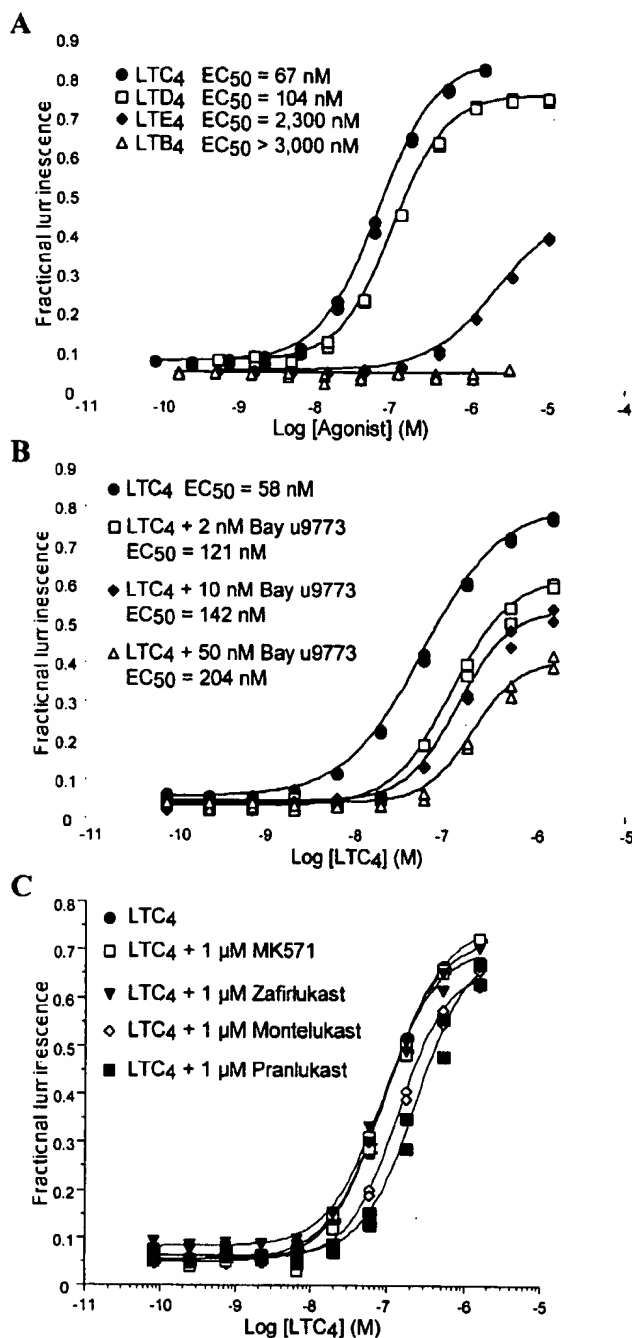


FIG. 3. Functional activation of HG57, the human CysLT₂ receptor expressed in HEK 293T cells using the aequorin assay. A, representative concentration response curves from HEK 293T cells co-transfected with pCDM-Aeq and human HG57 (CysLT₂)-pCR3.1 and challenged with leukotrienes. The fractional luminescence responses to LTC₄ (●), LTD₄ (□), LTE₄ (◆), and LTB₄ (△) are plotted as a function of their concentrations. B, antagonism of LTC₄-challenged HG57 (CysLT₂) receptor expressing HEK 293T cells by BAY u9773. A dose-response curve was first generated for LTC₄ alone (●), from 0.1 nM to 1 μM, followed by three additional dose-response curves in which cells were preincubated for 15 min in the presence of a fixed concentration of BAY u9773 of 2 nM (□), 10 nM (◆), or 50 nM (△) prior to challenge with increasing concentrations of LTC₄. C, antagonism of LTC₄-challenged HG57 (CysLT₂) receptor expressing HEK 293T cells by CysLT₂ receptor-selective antagonists. A dose-response curve was first generated for LTC₄ alone (●), from 0.1 nM to 1 μM, followed by one additional dose-response curve for each antagonist in which cells were preincubated for 15 min in the presence of a 1 μM concentration of either MK571 (□), montelukast (◇), zafirlukast (▼), or pranlukast (■) prior to challenge with increasing concentrations of LTC₄. For A, B, and C, EC₅₀ values shown on the graph were determined as described under "Experimental

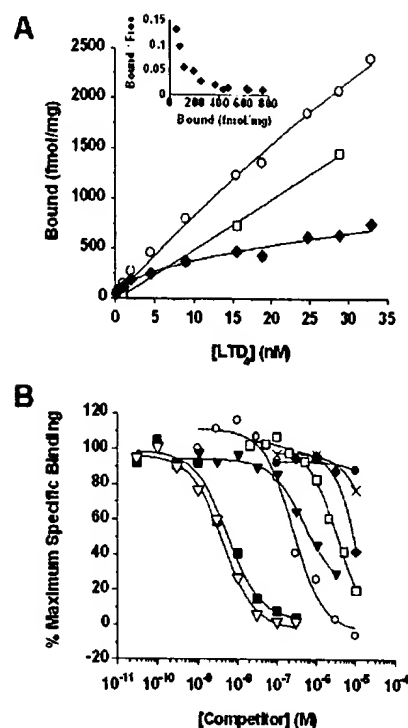


FIG. 4. Radioligand binding studies with COS-7 cell membranes expressing HG57, the human CysLT₂ receptor. A, saturation analysis of [³H]LTD₄ binding to HG57 (CysLT₂). Radioligand binding assay was performed in the presence of increasing amounts of [³H]LTD₄ supplemented with unlabeled ligand (final specific activity of [³H]LTD₄ = 26.8 Ci/mmol) over a range of concentration of 0.04–33 nM total LTD₄. The calculated [³H]LTD₄-specific binding saturation isotherm was analyzed by nonlinear transformation using Prism (GraphPad Software Inc.) to generate K_d and B_{max} values. A representative saturation curve from two independent experiments done in duplicate with two different membrane preparations is shown. Inset, Scatchard representation of the specific binding isotherm. ○, total; □, nonspecific; ◆, specific. B, equilibrium competition assays for [³H]LTD₄-specific binding to HG57 (CysLT₂) with leukotrienes and CysLT₂-specific antagonists. Representative titration curves from at least two independent experiments done in duplicate are shown. Competition curves were analyzed with a custom designed software using a nonlinear least-squares curve fitting routine based on a four-parameter logistic equation to determine half-maximal inhibitory concentration (IC₅₀) values. ▼, LTC₄; ■, LTD₄; ▼, LTE₄; ●, LTB₄; ○, BAY u9773; ×, montelukast; ◆, zafirlukast; □, pranlukast.

In one case (Fig. 4A), data analysis using Prism (GraphPad Software Inc.) revealed the presence of high and low affinity binding sites (K_d = 0.4 and 51 nM; B_{max} = 0.135 and 1.415 pmol/mg of membrane protein, respectively) as illustrated in the Scatchard representation of the deduced specific binding isotherm (Fig. 4A, inset). However, in the second case, saturation analysis showed the presence of a single population of binding sites (K_d = 4.8 nM; B_{max} = 0.338 pmol/mg of membrane protein). This difference is potentially due to variations between membrane preparations inherent in using a transient expression system. We are currently developing clonal cell lines expressing CysLT₂ that should permit a more accurate assessment of ligand:receptor affinity and receptor number. In equilibrium competition assays, the rank order of potency of leukotriene agonists to compete with [³H]LTD₄ for binding to the CysLT₂ receptor was LTD₄ = LTC₄ >> LTE₄ with no competition up to 10 μM by LTB₄ (Fig. 4B). CysLT₁ receptor antagonists were either weak (zafirlukast and pranlukast) or inactive

Procedures." Data are from a representative of three experiments, performed in duplicate.

FIG. 5. Northern blot analysis of human tissue RNA. RNA was prepared from human tissues as described under "Experimental Procedures" and hybridized with a 32 P-labeled fragment of the CysLT₂ receptor. X-ray films were exposed for 3 days for all tissues except brain, which was exposed to film for 7 days.

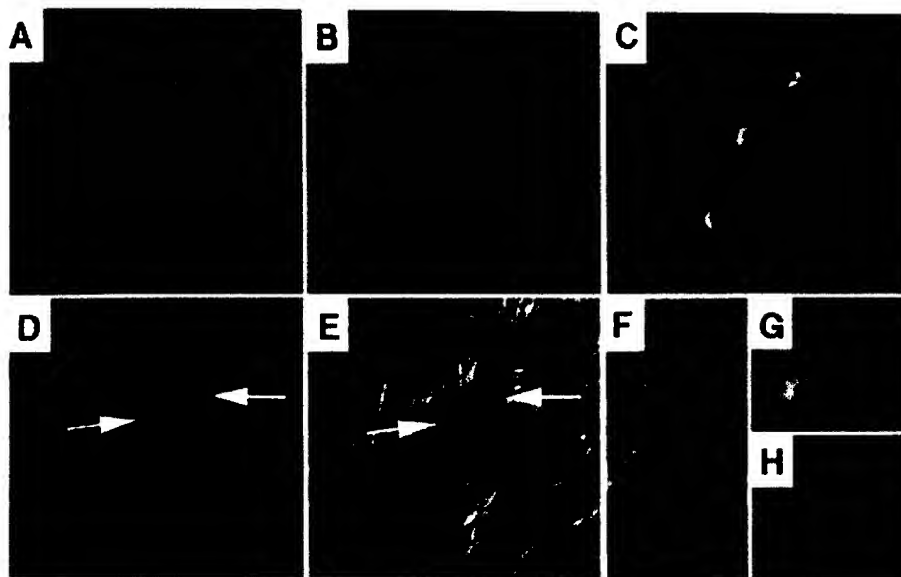
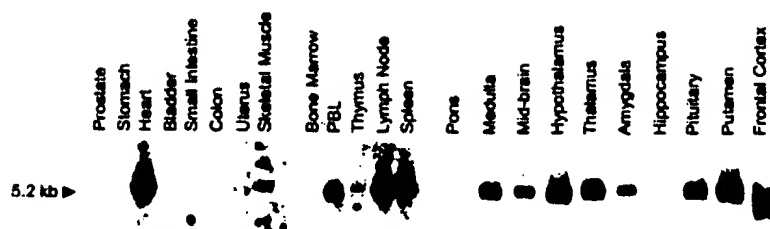
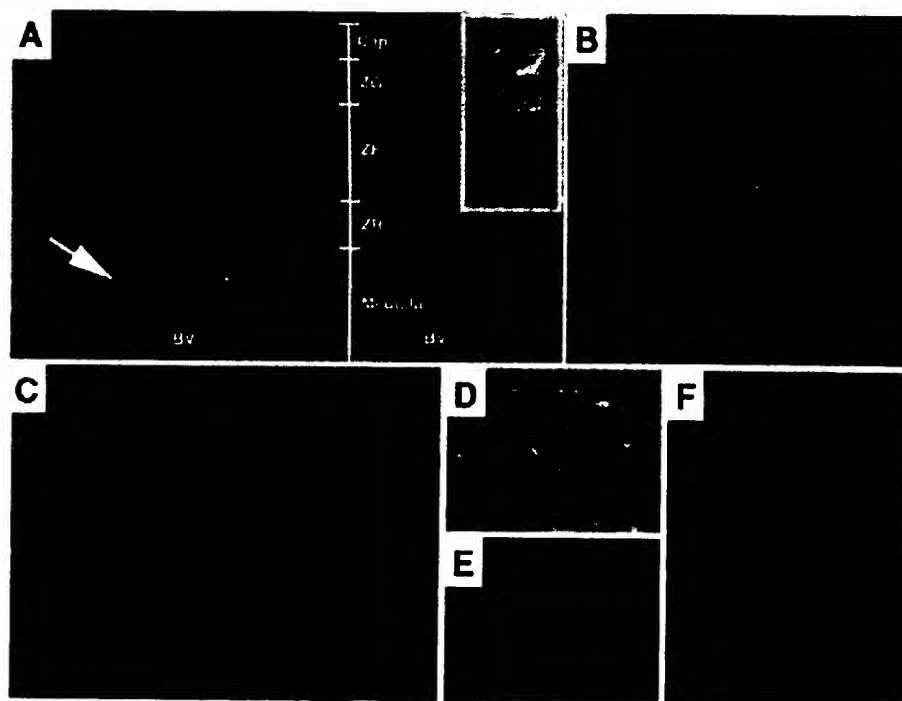


FIG. 6. *In situ* analyses of the human CysLT₂ receptor in lung and blood. A-E, *in situ* hybridization in human lung. A, antisense probe positive (Texas Red) in human lung interstitial cells; 4,6-diamidino-2-phenylindole blue nuclei. B, immunohistochemical staining (fluorescein isothiocyanate green positive signal) of LN5, a macrophage-specific antigen marker, on interstitial lung cells from A. C, yellow signal of combined detection of antisense (Texas Red) CysLT₂ receptor, and immunohistochemistry (fluorescein isothiocyanate green) LN5 macrophage marker on interstitial lung cells indicating colocalization. D, antisense probe on lung smooth muscle cells and abundant expression on adjacent macrophages. E, differential interference contrast image of section. F-H, *in situ* hybridization in human peripheral blood cells. F, antisense probe, Texas Red-positive in peripheral blood monocytes; teal blue pseudocolor shows nuclei. G, antisense probe in purified human eosinophils. H, sense control probe on the same population of eosinophils.

FIG. 7. *In situ* analyses of the human CysLT₂ receptor in adrenal medulla and heart. A, antisense probe in adrenal medulla (inset, high magnification reveals pheochromocyte). B, sense control probe in adrenal medulla. C, antisense probe in human heart. D, Nomarski image of heart showing typical morphology of a Purkinje conducting cell. E, antisense probe in heart Purkinje cell shown in D. F, sense control probe of section in C.



(montelukast) at competing for radiolabeled LTD₄ binding to the CysLT₂ receptor (Fig. 4B and Table I). However, full competition was observed with the dual CysLT₁/CysLT₂ antagonist BAY u9773 with an IC₅₀ value of 0.6 μ M (Fig. 4B and Table I). COS-7 cell membranes transfected with vector alone showed no specific LTD₄ binding, as previously reported (4).

Human Tissue RNA Northern Blot and *In Situ* Analyses—

The human CysLT₂ receptor was shown to be expressed in peripheral blood leukocytes, lymph node, spleen, heart, and several central nervous system regions (Fig. 5). In human RNA dot blot analyses, the receptor was also shown to be expressed in the adrenal gland (data not shown). The CysLTs are known to have potent contractile and inflammatory effects in the human lung, so we investigated the *in situ* expression of

TABLE I
Affinities of leukotrienes and CysLT antagonists for HG57, the human CysLT₂ receptor

Equilibrium competition assays were done as described under "Experimental Procedures" and in the legend to Fig. 4. Results are expressed as mean IC₅₀ ± S.E., where n = 3; otherwise, individual values are shown.

Competitor	IC ₅₀
	nM
LTC ₄	3.35 ± 0.53
LTD ₄	3.48 ± 1.01
LTE ₄	693 ± 14
LTB ₄	>10,000, >10,000
BAY u9773	597 ± 297
Pranlukast	3620 ± 755
Zafirlukast	7397 ± 871
Montelukast	>10,000, >10,000

CysLT₂ receptor mRNA in human lung (Fig. 6, A–E). In contrast to the CysLT₁ receptor, which was most highly expressed in human lung smooth muscle (4), the strongest lung expression of the CysLT₂ receptor was seen in interstitial macrophages (Fig. 6, A–C) with distinctly weaker expression in smooth muscle cells (Fig. 6, D and E). It was of interest that we noted particularly elevated expression of the CysLT₂ receptor in macrophages when in close proximity to smooth muscle cells (Fig. 6D). Peripheral blood leukocytes were another abundant source for CysLT₂ receptor transcripts (Fig. 5). We carried out *in situ* analysis of the CysLT₂ receptor in partly purified normal human peripheral blood monocytes and demonstrated abundant expression in greater than 20% of these cells (Fig. 6F). In addition, we found strong expression of the receptor in purified human eosinophils, a cell type in which we had previously shown expression of the CysLT₁ receptor (Ref. 4; Fig. 6, G and H).²

Tissues other than myeloid cell containing organs, such as the spleen and lymph nodes, that showed expression of the mRNA for the CysLT₂ receptor included regions of the central nervous system, the adrenal gland, and the heart. The synthesis and activities of cysteinyl leukotrienes in the brain of various animal species has been well documented (21–23). Studies on expression of both the CysLT₁ and CysLT₂ receptors in the central nervous system are under way in our laboratories. *In situ* analyses of the CysLT₂ receptor in adrenal gland showed expression in medullary pheochromocytes and ganglion cells (Fig. 7A). Pheochromocytes were identified by their morphology and granular nonhomogenous cytoplasm. CysLT₂ receptor expression was also noted in the zona reticularis adjacent to the corticomedullary junction. The expression of the CysLT₂ receptor in the adrenal gland was unexpected, and the biological roles of CysLTs in the adrenal medulla are unknown. In the human heart, CysLT₂ receptor expression was concentrated in Purkinje fiber cells, identified by their characteristic morphology as transitional cells with large nuclei in histologically inhomogenous areas, with few myofibrils (Fig. 7, C–E). Consistent with Purkinje cell identity, these CysLT₂ receptor-expressing cells co-expressed PGP9.5 (protein gene product 9.5) (data not shown). In contrast, the CysLT₁ receptor has not been detected in the heart (4, 5). A considerable pharmacological literature exists on the synthesis and actions of CysLTs in human vasculature (24–26). The restricted expression of CysLT₂ in the heart is intriguing.

The human CysLT₂ receptor was localized to chromosome 13q14 by radiation hybrid and somatic cell hybrid analyses. The marker closest to the CysLT₂ receptor gene was D13S153. This is of particular interest, since this marker has been associated with atopic asthma in both a United Kingdom population (27) and in a Japanese population study (28). We are

investigating the possible linkage of polymorphisms within the CysLT₂ receptor gene to the asthmatic phenotype. The recent deposition of a bacterial artificial chromosome in the high throughput genome sequencing division of GenBank™ (clone_id RP11–108P5) containing the precise HG57 (CysLT₂) receptor open reading frame reports a map position of 13q14.2–21.1, which is in agreement with our experimentally determined localization. The cloning and characterization of the human CysLT₂ receptor will allow a detailed molecular characterization of its activation in pulmonary, hematologic, cardiovascular, endocrine, and central nervous system tissues.

Acknowledgments—We acknowledge Drs. D. Pettibone, R. Gould, and A. W. Ford-Hutchinson for support of the Merck Research Laboratories orphan GPCR project. We also thank Kristine Waldburger and Bruce Daugherty for the preparation of human peripheral blood monocytes and eosinophils, Sam Lee for graphics, and Sandy Camburn for secretarial support.

REFERENCES

- Samuelsson, B. (1983) *Science* **220**, 568–575
- Coleman, R. A., Eglen, R. M., Jones, R. L., Narumiyi, S., Shimizu, T., Smith, W. L., Dahlen, S. E., Drazen, J. M., Gardiner, P. J., Jackson, W. T., Jones, T. R., Krell, R. D., and Nicosia, S. (1995) *Adv. Prostaglandin Thromboxane Leukotriene Res.* **23**, 283–285
- Metters, K. M. (1995) *J. Lipid Mediat. Cell Signal.* **12**, 413–427
- Lynch, K. R., O'Neill, G. P., Liu, Q., Im, D.-S., Sawyer, N., Metters, K. M., Coulombe, N., Abramovitz, M., Figueroa, D. J., Zeng, Z., Connolly, B. M., Bai, C., Austin, C. P., Chateaufneuf, A., Stocco, R., Greig, G. M., Kargman, S., Hooks, S. B., Hosfield, E., Williams, D. L., Jr., Ford-Hutchinson, A. W., Caskey, C. T., and Evans, J. F. (1999) *Nature* **399**, 789–793
- Sarau, H. M., Ames, R. S., Chambers, J., Ellis, C., Elashourbagy, N., Foley, J. J., Schmidt, D. B., Muccitelli, R. M., Jenkins, O., Murdock, P. R., Herrity, N. C., Halsey, W., Sathe, G., Muir, A. L., Nuthulaganti, P., Dytko, G. M., Buckley, P. T., Wilson, S., Bergsma, D. J., and Hay, D. W. (1999) *Mol. Pharmacol.* **58**, 657–663
- Reiss, T. F., Chervinsky, P., Dockhorn, R. J., Shingo, S., Seidenberg, B., and Edwards, T. B. (1998) *Arch. Intern. Med.* **158**, 213–220
- Suissa, S., Dennis, R., Ernst, P., Sheehy, O., and Wood-Dauphinee, S. (1997) *Ann. Int. Med.* **126**, 177–183
- Grossman, J., Faiferman, I., Dubb, J. W., Tomsom, D. J., Busse, W., Bronaky, E., Montanaro, A., Southern, L., and Tinkelman, D. (1997) *J. Asthma* **34**, 321–328
- Tudhope, S. R., Cuthbert, N. J., Abram, T. S., Jennings, M. A., Maxey, R. J., Thompson, A. M., Norman, P., Gardiner, P. J. (1994) *Eur. J. Pharmacol.* **264**, 317–323
- Labat, C., Ortiz, J. L., Norel, X., Gorenne, I., Verley, J., Abram, T. S., Cuthbert, N. J., Tudhope, S. R., Norman, P., Gardiner, P., and Brink, C. (1992) *J. Pharmacol. Exp. Ther.* **263**, 800–805
- Walch, L., Norel, X., Gascard, J. P., and Brink, C. (1999) *J. Physiol. Pharmacol.* **50**, 567–573
- Retief, J., Lynch, K. R., and Pearson, W. R. (1999) *Genome Res.* **9**, 373–382
- O'Dowd, B. F., Nguyen, T., Lynch, K. R., Kolakowski, L. F., Thompson, M., Cheng, R., Marchese, A., Ng, G., Heng, H. H. Q., and George, S. R. (1996) *FEBS Lett.* **394**, 325–329
- Durieux, M. E., Carlisle, S. J., Salafra, M. N., and Lynch, K. R. (1993) *Am. J. Physiol.* **264**, C1360–C1364
- Hooks, S. B., Ragan, S. P., Hopper, D. W., Honemann, C. W., Durieux, M. E., MacDonald, T. L., and Lynch, K. R. (1998) *Mol. Pharmacol.* **53**, 188–194
- Unguin, M. D., Singh, L. M., Stocco, R., Sas, D. E., and Abramovitz, M. (1999) *Anal. Biochem.* **272**, 34–42
- Frey, E. A., Nicholson, D. W., Metters, K. M. (1993) *Eur. J. Pharmacol.* **244**, 239–250
- Marchese, A., Docherty, J. M., Nguyen, T., Heiber, M., Cheng, R., Heng, H. H., Tsui, L. C., Shi, X., George, S. R., and O'Dowd, B. F. (1994) *Genomics* **23**, 609–618
- Chomezynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159
- Yokomizo, T., Izumi, T., Chang, K., Takauwa, Y., Shimizu, T. (1997) *Nature* **387**, 620–624
- Lindgren, J. A., Hokfelt, T., Dahlen, S.-E., Patrono, C., and Samuelsson, B. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6212–6216
- Hulting, A. L., Lindgren, J. A., Hokfelt, T., Eneroth, P., Werner, S., Patrono, C., and Samuelsson, B. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3834–3838
- Winking, M., Deinsberger, W., Joedicke, A., and Boeker, D. K. (1998) *Cerebrovasc. Dis.* **8**, 318–326
- Allen, S. P., Chester, A. H., Piper, P. J., Sampson, A. P., Akl, E. S., and Yacoub, M. H. (1992) *Br. J. Clin. Pharmacol.* **34**, 409–414
- Vigorito, C., Giordano, A., Cirillo, R., Genovese, A., Rengo, F., and Marone, G. (1997) *Int. J. Clin. Lab. Res.* **27**, 178–184
- Folco, G., Rossoni, G., Buccellati, C., Berti, F., Maclof, J., and Sala, A. (2000) *Am. J. Respir. Crit. Care Med.* **161**, S112–S116
- Cookson, W. (1999) *Nature* **402**, B5–B11
- Kimura, K., Noguchi, E., Shibasaki, M., Arinami, T., Yokokouchi, Y., Takeda, K., Yamakawa-Kobayashi, K., Matsui, A., and Hamaguchi, H. (1999) *Hum. Mol. Genet.* **8**, 1487–1490

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.